DOCUMENT RESUME

ED 254 921 EA 017 584

AUTHOR Wholeben, Brent Edward
TITLE Validating Multivariate Decision Modeling for Educational Planning.
PUB DATE Oct 84
NOTE 80p.; Paper presented at the Annual Meeting of the International Society for Educational Planning (New Orleans, LA, October 17-19, 1984).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Comparative Analysis; Computer Simulation; *Decision Making; Elementary Secondary Education; *Evaluation Methods; Mathematical Models; *Models; *Multivariate Analysis; Problem Solving; Validity
IDENTIFIERS Decision Support Systems; *Multiple Alternatives Model

ABSTRACT

Complex decision-making environments often demand that a range of alternative strategies be identified and compared for their shared value in achieving desired ends. Simple parametric statistical procedures such as analysis of variance and discriminant function analysis can be used in such situations to compare (using selected criteria) the value of one set of decisions over another. The Multiple Alternatives Model (MAM), formulated by the author in 1978, is described as a system of linear equalities and inequalities, whose dependent variables (the "unknowns") are the various alternatives being evaluated for possible entry into the final solution set, and whose coefficients are criterion values that constrain the comparison and contrast of these alternatives during evaluation. This paper discusses the use of this model in various decision-making contexts. Topics discussed include validation of criteria-based decision-making, structure of alternatives in decision-making, content and process considerations, statistical utilities for decision model validation, tools and mechanics of MAM decision validation, and special considerations in hierarchical model validation. A 16-item list of selected readings and references is included. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

BEST COPY

ED254921

VALIDATING MULTIVARIATE DECISION MODELING FOR EDUCATIONAL PLANNING

Brent Edward Wholeben, Ph.D., E.M.T.
Associate Professor of Educational Administration

Department of Educational Leadership and Counseling
The University of Texas at El Paso (El Paso, Texas 79968)

Presented to the 1984 Annual Meeting of the
International Society of Educational Planners (ISEP)
New Orleans, Louisiana (October 1984)

ABSTRACT

Standard multivariate, parametric statistical procedures can be easily employed when educational planners wish to validate potential decisions against stated criteria.

SUMMARY

Complex decisioning environments which demand that a range of potential decision alternatives (alternate strategies) be identified, and then summarily compared for their shared value in achieving desired ends, often allude a useful technique through which the final decision might be evaluated against the other alternatives which were not selected as the 'decision'. Simple parametric statistical procedures as analysis of variance and discriminant function analysis can be used in such situations to allow a comparision against selected criteria of the comparative value in one set of decisions over another. With the wide availability of computer-based statistical analysis packages existing today, the deployment of statistical techniques is yet an oft-forgotten possibility.

EA 017 584

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

- ALL RIGHTS RESERVED -

(Printed in the United States of America)

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Brent Edward Wholeben
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

D E D I C A T I O N

This project is dedicated to all program evaluators, policy analysts and key staff advisors who perform decisionmaking based upon one set of criteria and announce their decisions with references to another set of criteria -- incessantly modifying their decisions because they can not provide the supporting rationale and information based articulation in the face of public protest and inquiry -- with familial assurances to the media that the sensitivity and complexity of the situation prevent more reasonable explanations.

and

To Doug and Don ... who, throughout a highly volatile and sensitive era of individual human rights versus numbers in a computer, sought the truth and accepted my findings though the hangman was outside the door.

© 1984 by Brent Edward Wholeben, Ph.D., E.M.T.

No part of this monograph may be reproduced in any form, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the author.

- All Rights Reserved -

BEST COPY

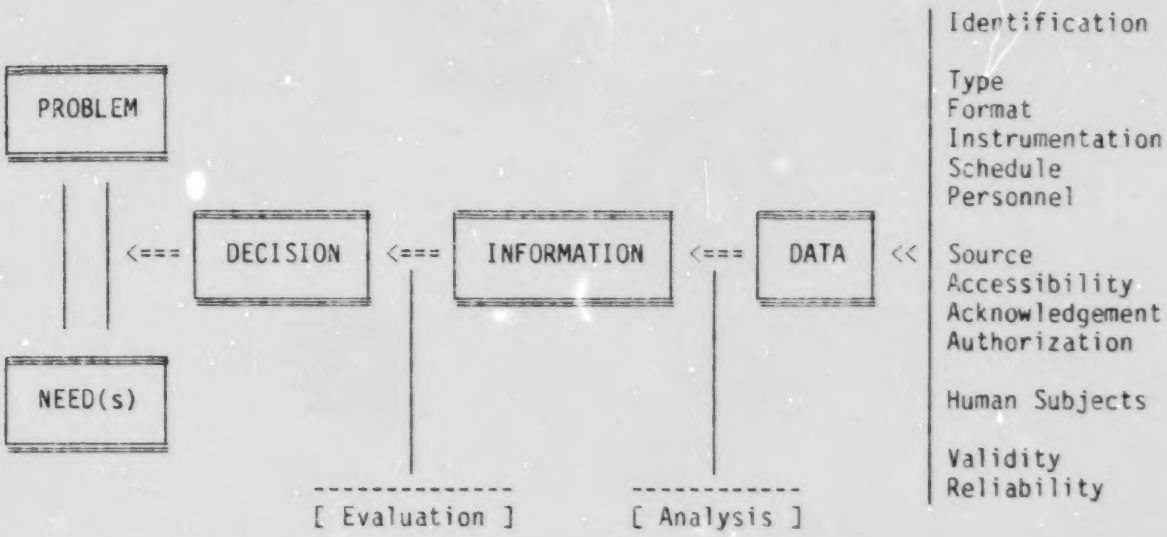
INTRODUCTION

The wide-spread availability of computers during the past few years, and the increased facility and comfortability which this availability has fostered, might reasonably have predicted an increased deployment of computers for educational planning and evaluation. During the past ten years, education has been faced with some of its most complex and intricate problems to date -- declining enrollments, limited fiscal revenues, decreasing student learning, racial and cultural desegregations, urban flight, transient student populations -- but little of the potential of the computer for helping to find solutions to these problems has been forthcoming.

Indeed, the computer has often been viewed as an inappropriate tool for decisionmaking because of its total reliance on validly defined and reliably measured data with which to produce relevant and timely information for decisionmaking, the validity and reliability of data often being itself in question. The use of data for decisionmaking, especially data which has been quantified, is also highly suspect by many decisionmakers. Numerical data, much like many people's reluctance and fear of mathematics, is seen as something to be feared, untrustworthy, and therefore to be avoided. These views, of course, are most unfortunate, since all elements of data -- quantitative or qualitative -- with assurances of appropriate validity and adequate reliability, are strong allies of the decisionmaking process.

Decisionmaking typically involves the analysis of data regarding alternative courses of action, a subsequent comparision and contrast of these alternatives, and finally the choice of that alternative (or those alternatives) which seems best to fit the problematic situation at hand. This entire process of alternatives' comparision and selection must be based upon available data.

BEST COPY



Many elements are involved in the generation of data which is finally made available to the decisionmaking process, including the identification of what data is needed based upon the problem situation, technical aspects of its type, format, instrumentation and source, the protocol for securing accessibility to the data, the parallel human subjects review standards in force, and finally those elements of validity and reliability which must be taken into account before the decisionmaker can place adequate faith in the data itself. Data is then analyzed and reported to the decisionmaker is an appropriate form so that the data provides information to the decisionmaking process. The information so provided is evaluated by the decisionmaker -- that is, alternative courses of action are compared and contrasted based upon their individual value and benefit -- and a decision selected as the 'best approach' to the problem at hand based upon the information available from the data gathered. A seemingly simple and obvious approach, to be sure, but one which is often ignored because of time, reluctance or ignorance.

Furthermore, even if a data based and information referenced approach to decisionmaking is taken by the decisionmaker, the results of the decisionmaking process -- that alternative course of action selected for implementation -- is seldom post hoc analyzed and verified for its authenticity and utility, viz. validated. If care and discretion are employed during the data identification

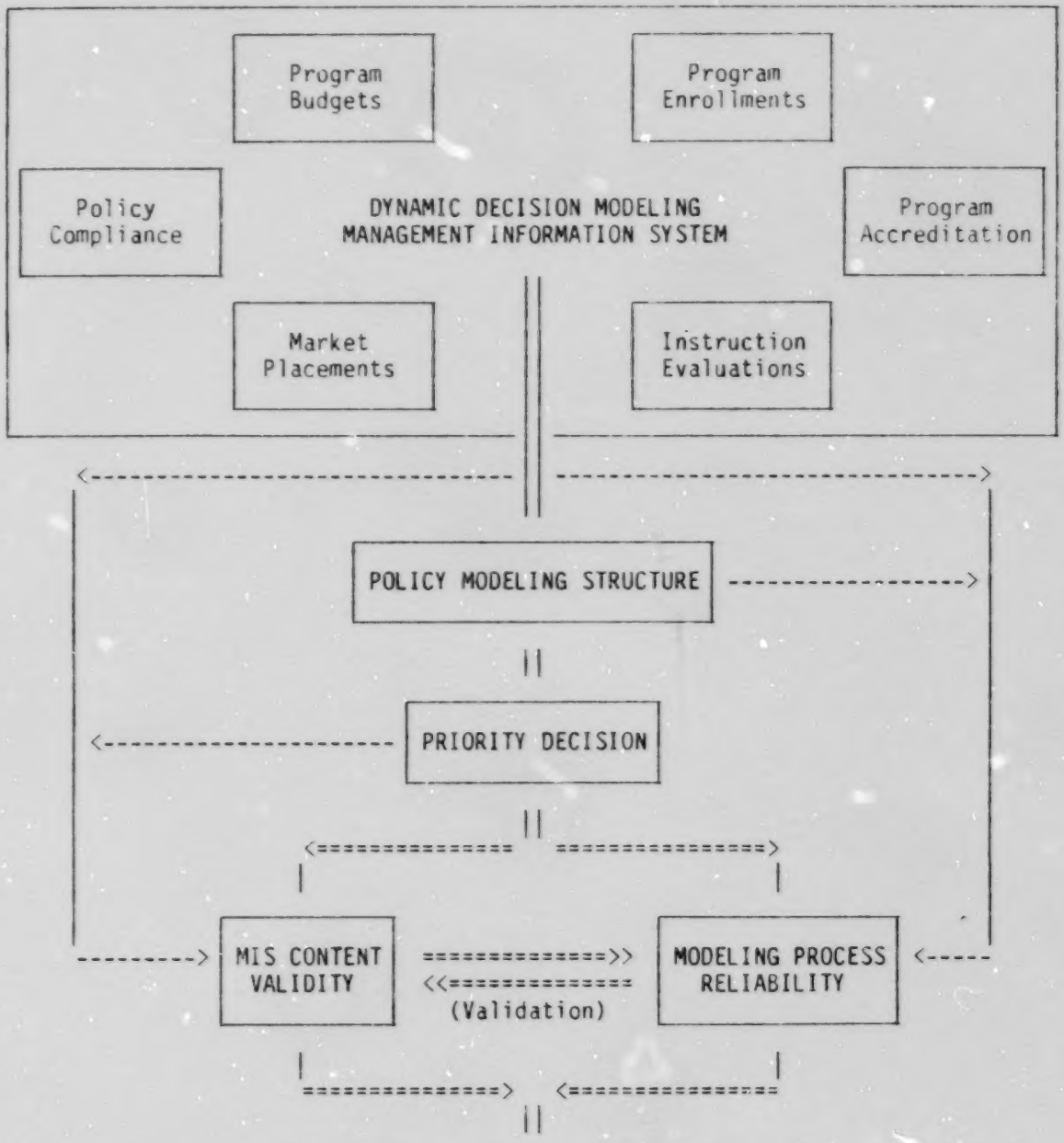
BEST COPY

and gathering episodes, if information is generated by suitable means, and that information applied in a reasonable manner, is a post hoc validation of the data-information-decision process really necessary?

No problem exists within a vacuum, and therefore the specific course of action which will be deployed to solve the problem must also be viewed as a product of a dynamic, not static, environment. Data available to a decision must attest to its validity and reliability -- validity which is often arguable, but reliability which is always variable. Data elements are constantly changing references, and their measurement at any point in time mere samplings of their actual value or statement of product. The decisionmaking process must therefore be equally variable, or dynamic, and moreover capable of controlling for the variability in information available to the alternatives comparison and contrast processes. Information in this sense is viewed as criteria through which the decision will be identified and valued, and subsequently indicative of the final solution designated by the decisionmaker as the best approach possible.

This idea of dynamic information utility is schematically represented in the following diagram.

BEST COPY



VALIDATION OF POLICY MODELING TECHNIQUE, and
SUPPORTING MANAGEMENT INFORMATION SYSTEMS

Information available to the decisionmaking process on a timely basis can be defined as a dynamically oriented management information system. The management

BEST COPY

information system, or MIS, gathers data from its interactive environment, and directs that data in the form of information into a policy modeling structure. The policy modeling structure is defined as the use of data based information for identifying, evaluating and selecting likely courses of action which are in compliance with established policy. A priority decision is then made which bests fits the information provided, the problem encountered, and the policy of the organization in question. What remains is the post hoc assessment of this final decision (or decisions) in terms of its (1) compliance with the criteria used in the decisionmaking process (i.e. the MIS content) and its (2) adequacy and appropriateness based upon the organizational policy modeled.

It is the purpose of this paper to explore why such a post hoc assessment should be implemented, the tools and techniques associated with implementation, and the utility of various parametric and nonparametric statistical procedures for evaluting the adequacy and appropriateness of the decisions selected.

DECISIONMAKING THROUGH MODELING SIMULATIONS

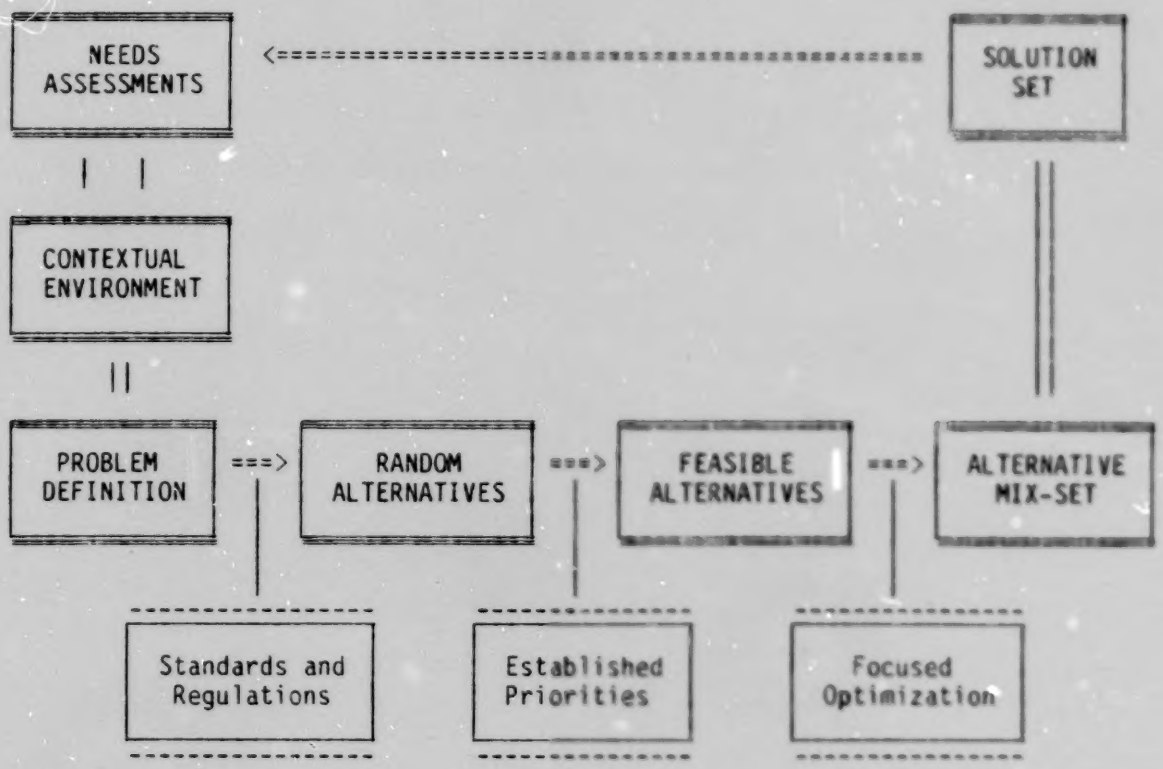
As a general rule, decisionmaking is the adequate and appropriate evaluation of all alternative courses of action available to the decisionmaking process based upon the problematic situation defined, and the subsequent selection of some one or more alternatives based upon its (their) relative value and benefit. Each alternative course of action is considered to be reasonable if it displays some value in remediating the problem situation. If, in fact, the alternative meets or exceeds the standards of the decisionmaking process -- limits placed upon the values of the criteria to signify minimum positive effect necessary and maximum negative impact allowable -- then the alternative is termed feasible.

Typically, the reasonableness of any alternative is defined by measured characteristics and/or attributes of the alternative which match similarly measured needs of the environment in which the problem resides. This ability to use criterion based measures to define the problem, its environment and the alternative(s) available to remediate the problem is called <u>modeling</u>; that is, a model of the decisionmaking setting is constructed using data referenced 'points

BEST COPY

of purpose' (POPs) known as criteria. The further ability to structure this model in such a way as to predict the impact of the various alternatives on the problem, and thereby select that alternative (or alternatives) which best fits the problem modeled, is called simulation; that is, to forecast (simulate) or predict likely effects based upon measured attributes of the modeled environment, its problem(s), and the alternatives available to provide remediation.

Such a model might generally be viewed as follows:



Note how the standards and regulations of the organization allow an initial declaration of alternative courses of remedial action, called random alternatives, which might provide relief to the system modeled. Established priorities of such a system, however, typically refine random alternatives into a smaller set of actions which are preferred, based upon these established priorities which are normally dynamic in terms of an ever changing environment, called

BEST COPY

feasible alternatives, viz. feasible solutions based upon current priorities. From this set of feasible alternatives, one or more optimal courses of actions are selected for their 'best-fit' characteristics with regard to the problem model. These actions will then be executed as the solution sought.

Consider the problem situation of district wide, declining student enrollments, where the policy to be modeled is that of targeting schools for closure based upon each school's individual characteristics or attributes. The purpose of such an evaluation model would be to define, develop and display a formulation which has significant potential for exploring the myriad of decisions available to the school district officer when faced with the need to evaluate each of the district's campus sites for potential future closure. An appropriate mission statement for such a model could be formulated as:

> To design a system of alternative school closure options, accounting for differential levels of operational discontinuance or continuance for each school site evaluated, based upon likely scenarios for the future needs and demands of the school district and its community, which could be compared and evaluated against established policy standards, and instructional and administrative requirements, and which would result in the selection of <u>which</u> schools would be closed or allowed to remain open, and the <u>level</u> of that operational discontinuance or continuance, respectively.

Such a school closure model seeks to determine (1) the number of schools which should be targeted for closure, (2) the disposition of those closed schools, and (3) the operating level of those schools remaining open. This modeling technique allows the decisionmaker to utilize all available data and information for simulating the impact to the school district which alternative closure decisions per each individual site might provide. This forecasting strategy includes the ability to control for interrelationships between discrete pairs of programs in contrasted schools, as well as among the collective impact of groups of programs with their individual levels of recommended resource support. Finally, this modeling strategy allows the decisionmaker to vary the initial controls placed upon the model's mathematical structure, thereby simulating alternative decisional solutions based upon an ever fluid and transactional environment.

BEST COPY

Another example of a decison model centers around the need to determine resource allocations and distributions based upon scant resource availability, especially in the case of budgetary reductions or rollbacks. Here, the decisionmaker seeks to define, develop and display a mathematical modeling formulation which explores the myriad of decisions available to the educational administrator when faced with the need to evaluate discrete educational program budgets (instructional, service and administrative) for possible future funding variances. An appropriate mission statement of such a model for educational fiscal crisis budgeting at the university level might appear as:

> To design a system of alternative budgeting alternatives, based upon likely scenarios for the future needs and demands of the university, and its individual colleges and/or schools which could be compared and evaluated against established policy standards, and instructional and administrative requirements, and which would result in the selection of which program budgets will be reduced, and by how much.

To better understand the 'modeling' perspective for this mission of graduated budget reductions (identification and assignment), several goals must be accomplished (or satisfied) which support the successful accomplishment of this mission. For example:

> Prepare, for each program budget, the types and amounts of budgeting reductions which could be reasonably applied to the program budget, and in addition, prepare a narrative which would best describe the program after each type/amount of reduction;
>
> Define, across all program budgets, those measures of criterion reference which would best describe the resource needs, program contributions to university and individual college and/or school goals, and internal/external perceptions of the programs whose budgets are being evaluated;
>
> Select, for each criterion reference defined in Goal 2, the maximum (in the case of negative impact) and minimum (in the case of positive impact) effects to the university and individual colleges and/or schools which are sought based upon the types of programs reduced, and the amounts of each of their reductions;

BEST COPY

Assure, for each program being evaluated, that one and only one level (or type) of reduction (if any) is selected;

Determine which of the criterion comparison references will be utilized to guide the final decision of which budgets will be selected for reduction, and the amount of that reduction, in the event that multiple university program reduction strategies become available to the university president, college deans, and school directors and/or department chairpersons;

Validate, for each of the criterion references utilized for determining the final types and amounts of program reductions, the extent to which each of these criterion references significantly determined the final selection of those programs for reduction, and the amount(s) of those reductions; and,

Correlate, in terms of the programs chosen for reduction and the weights associated with each of these reductions, the final overall reduction strategy of the university and the accuracy with which the university and individual colleges and/or schools can honestly trust the results of that strategy.

Like the mission related to school closure modeling, the decisionmaker could utilize all available data and information for simulating the impact to the university which alternative funding levels per each individual program might provide.

We are especially interested, though, in the last two supporting goals stated above. Both of these statements speak to the need to validate the solution ultimately conceived by the modeling formulation, rather than simply accept the final alternative selected as both valid and reliable for the problem at hand. This would be equally true for the school closure model. The failure to perform such a post hoc validation of modeling results does not imply faith in the model, only an unfortunate disinterest in the influence which the criteria played in the determination of the final solution, and the ability of the model to control this influence both adequately and appropriately.

We have heretofore referred to the modeling of decisionmaking as a technique which is (1) information based, (2) alternatives formulated, and (3) impact referenced. Such a model was formulated in 1978 by the author, called the Multiple Alternatives Model™ or MAM™, initially designed to evaluate a school

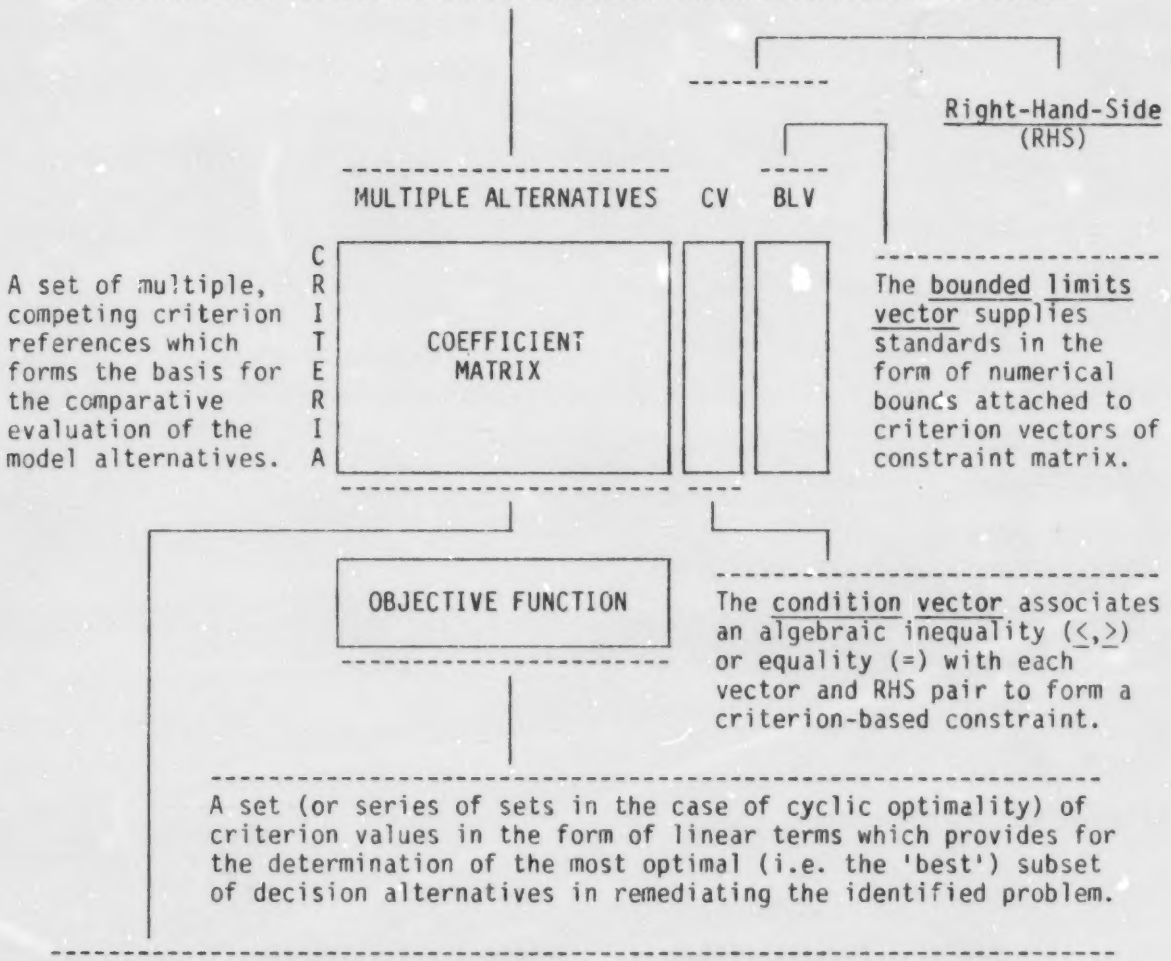
BEST COPY

closure decision facing a major metropolitan school district (Wholeben, 1980). Since then, similarly constructed models have been formulated to address such issues as budget reductions, district consolidations and grade-level reconfigurations by building (Wholeben, 1982; 1983). To better understand the techniques explored in the remainder of this paper concerning the validation of potential decisions, and the models which contributed to their identification, we will briefly viewed the MAM™ framework at this time.

The Multiple Alternatives Model™ is represented as a system of linear inequalities and equalities, whose dependent variables (the "unknowns") are the various alternatives being evaluated for possible entry into the final solution set, and whose coefficients are criterion values which constrain (viz. guide) the comparison and constrast of these alternatives during evaluation. The values of the dependent variables (viz. the decision alternatives) take on the value of '1' if selected as part of the solution, '0' otherwise. The MAM™ system might be viewed as follows:

BEST COPY

A collectively exhaustive set of mutually exclusive decisions which have been designated as comprehensively reasonable reactions to the problem situation identified; and from which a subset of decision alternatives may be drawn to construct the final solution set.


Right-Hand-Side
(RHS)
MULTIPLE ALTERNATIVES
CV
BLV
A set of multiple, competing criterion references which forms the basis for the comparative evaluation of the model alternatives.
CRITERIA
COEFFICIENT MATRIX
The bounded limits vector supplies standards in the form of numerical bounds attached to criterion vectors of constraint matrix.
OBJECTIVE FUNCTION
The condition vector associates an algebraic inequality (≤,≥) or equality (=) with each vector and RHS pair to form a criterion-based constraint.
A set (or series of sets in the case of cyclic optimality) of criterion values in the form of linear terms which provides for the determination of the most optimal (i.e. the 'best') subset of decision alternatives in remediating the identified problem.
The vector portion (coefficients of the left-hand-side) which denote for each of the modeled alternatives (viz., the dependent variables), the value of each particular criterion's measure or influence (viz., the independent variables) related to that same particular alternative. Thus, the coefficients of each term in the linear inequalities represent the specific criterion measure for that particular term (i.e. alternative).

For those readers interested in the technical characteristics of the model itself, you are directed to the technical papers, Modeling Elementary and Secondary School Closures Through MAM™ (1984), and Reducing Higher Education

BEST COPY

Budgets Through Multiple Alternatives Modeling™ (1984), each of which are available by writing directly to the author.

Basically albeit briefly the criteria which will be utilized to compare and contrast the relative benefit of each alternative are entered into the coefficient matrix as coefficients of the linear inequalities or equalities. These inequalities and equalities form the basis for mathematically constraining the selection of a subset of alternatives as the decision solution set, thereby allowing the criteria to compete for influence during the comparison and contrasting stage of model execution. This competition results in (normally) several subsets of alternatives being identified (e.g. schools to close) which 'fit the constraints of the model', each of these subsets being declared collections of feasible alternatives. The objective function is utilized to choose which of these subsets 'best approximates' the needs of the organization, that single subset then defined as the optimal solution or decision subset.

What the reader will recognize, but maybe not be totally comfortable with, is the strictly mathematical nature of the MAM™ decisioning system, and its reliance upon the conventions of number theory and arithmetic operations. How would a person validate such a solution as might result from a successful incorporation of this framework during decisional evaluation? How could the model be further analyzed to discern its ability to forecast the effects of later decisionmaking in this context? And finally, how could such knowledge allow the decisionmaker to validate other decisions which are not generated through a MAM™ oriented approach? We will explore the answers to these and other questions in the remaining sections of this paper.

VALIDATION OF CRITERIA BASED DECISIONMAKING

Validation means the determination of validity, or an assessment regarding the extent to which you accomplished subsequently what you set out to do initially. For decisionmaking based upon prespecified criterion standards, validation seeks to assess (1) the extent to which criteria influenced the final decision selected and (2) the role of the decisionmaking technique in intro-

BEST COPY

ducing the criteria into the comparison and contrasting process of the alternative courses of action. Validation of decisionmaking should therefore involves approaches which are oriented towards both content (criteria used to value each of the alternative courses of action) and process (techniques which employed the criteria to compare and contrast the alternatives). In this paper, we view the validation of decisionmaking as addressing both content and process considerations.

Validation obviously values the need for responsibility and accountability in the decisionmaking process. In addition, the search for the degree of validity in a decision provides further assurances that, at least to the degree that any decision can be considered the 'right' or 'best' decision, the course of action selected from all available alternatives matches the criteria imposed on that selection (viz. decisionmaking) process.

Validation in decisionmaking is parallel to the validation of educational practice and instructional programs which is commonly accepted in education today. Validation simply seeks to determine how a finished product matches the mission, goals and objectives which were stated during the initial planning, design and development of a program. The extent to which a developed program (1) remediates the needs which inspired the program, (2) follows the policy and practice guidelines instituted by the program, and (3) complements other related goals in the organization, all speak to the assessed validation of the program. Decisionmaking, as a policy and as a practice, can also be viewed as a program, and therefore can be assessed for its 'validity' much the same way as a program is validated after implementation.

The use of careful validation practices in criteria based decisionmaking also allows a somewhat research-oriented approach to later decisionmaking requirements. Since potential alternative courses of action are valued as decisions based upon the influence of the criteria used to compare and contrast these alternatives, the competitive confluence of all criteria interacts to identify the 'best possible alternative' based upon this criterion competition.

BEST COPY

STRUCTURE OF ALTERNATIVES IN DECISIONMAKING

If the reader accepts the notion of decisionmaking as the selection of an appropriate course (or courses) of action from a range of alternatives, then it necessarily follows that the structure of these 'alternatives' is important and non-trivial. In this section, we will briefly explore the attributes of structure for alternative actions which will be compared and contrasted based upon their criterion-measured value for remediating an identified problem situation.

First, a distinction must be made between what is called the policy alternative and what is referred to as a policy implementation alternative. Policy alternatives are generally considered to be global statements of general mission or purpose, typically providing a direction to later specific program development and problem remediation. For example, declining enrollments in elementary and secondary education might require the following policy alternatives be compared and evaluated for their perceived reasonableness as a solution: close schools, consolidate school districts, redistribute school grade-level configurations, continue operations as currently implemented, etc.. Once a policy alternative has been adopted, however, the specifics of implementing that policy must be ordained and executed, requiring a second comparison and contrasting of alternatives, called implementation alternatives. For example, implementation alternatives related to a policy of 'school closures' during periods of student enrollment declines might include: closure and sale, closure and lease, closure and mothball, as well as what operational variances might occur in schools which remain open, e.g. functional partitioning and sub-site lease.

Implementation alternatives, viz. the decision alternatives, will fall into one of three categories: singular, multiple or hierarchical. Singular decision alternatives refer to situations when one and only one alternative course of action will compose a decision. Such a situation closely parallels the comparision and contrasting of organizational and program policies. Multiple decisioning situations allow a final solution to incorporate more than just a single alternative course of action in the final decision. An example is the evaluation of several school sites for the purpose of targeting each of the campuses for either closure or operational continuance. Since more than one campus

BEST COPY

can, and probably will be targeted for closure in most situations, the decision is considered multiple in nature. The hierarchical modeling of decisionmaking is also multiple in format, but attaches sub-levels of decisionmaking to each of the alternatives being considered as solutions. An example is the school closure model previously mentioned in which levels of closure (sale, lease and mothball) are attached to the initial decision to close a particular site.

The reader might now recognize two additional attributes of decision alternatives which must be incorporated into their structure: mutual exclusiveness and collective exhaustiveness. To be mutually exclusive, each decision alternative must have the ability to represent a course of action totally divorced from any of the other alternatives being considered. Therefore, the decision model should have the opportunity to choose alternative A or alternative B without the necessary proviso that choosing one automatically selects the other. As collectively exhaustive, the alternatives available for evaluation must (or at least, should) represent all alternative courses of action available to the decisionmaker considering the situation being modeled. To restrict any potential alternative from consideration by the model would preclude the influence of that alternative in the model's later decision to select or not select that alternative as part of the decisional solution set. Furthermore, the preclusion of an alternative, and that alternative's criterion values, from model evaluation also effectively excludes the model understanding the impact to the organization in that alternative's potential existence.

To further illustrate each of these attributes, we will examine the school closure model discussed in Modeling Elementary and Secondary School Closures Through MAM™ (Wholeben, 1984). The structure for each alternative may be demonstrated as follows:

BEST COPY

Alternative-j (j=1,...,n) ==> School Campus/Site				
Option-1	Option-2	Option-3	Option-4	Option-5

where: Option-1 = target the site for future closure, and prepare the buildings for sale and future relinquishment of district control;

Option-2 = target the site for future closure, and prepare the buildings for lease under district control;

Option-3 = target the site for future closure, and prepare the buildings for 'mothballing';

Option-4 = target the site for functional partitioning with partial physical units of the campus for sub-site lease; and,

lease; and,

Option-5 = target the site for continued full operation as an instructional services center.

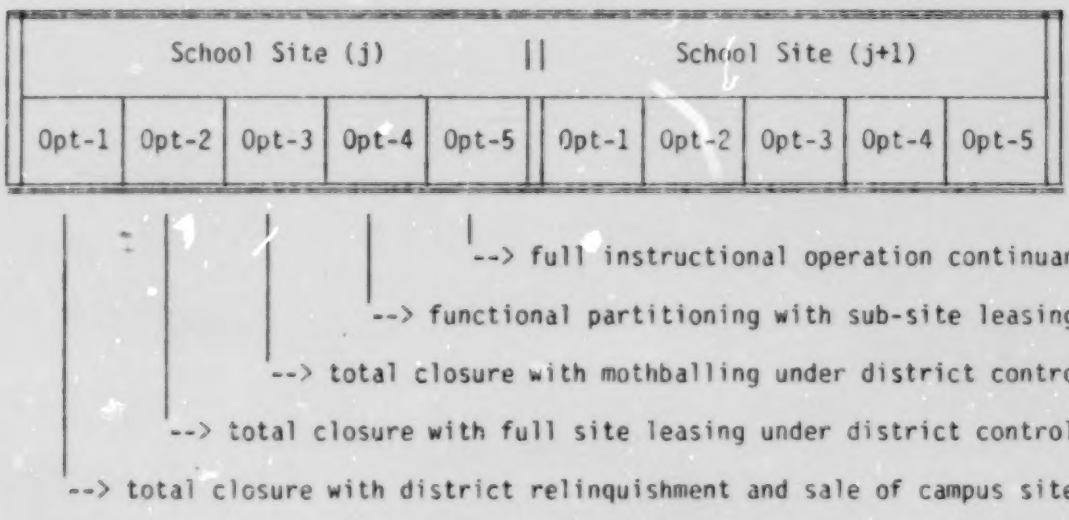
In this hierarchical school closure model, the decision to continue or discontinue any particular site's instructional operation(s) must be paralleled by the level of that continuance or discontinuance, respectively. In the decision model, a site that might be closed if it could be leased may be precluded from closure if criteria indicate that leasing is a remote possibility. Thus the interaction in a hierarchical decision model is geometrically increased from that which is found in the simple multiple alternatives setting (i.e. to close or not to close, without any reference to the level of closure).

To meet the requirements of mutual exclusiveness and collective exhaustiveness, the hierarchical school closure model must present alternatives for each

BEST COPY

school campus which are separate and distinct from other operational alternatives available to that same school, and must represent all available options to be considered for that particular school. Each of the five above options will be assessed across all schools, further classifying this hierarchical model as fixed-level.

The importance in defining the structure of these alternatives is better understood when different higher-order alternatives (i.e. the schools) are also contrasted along with the lower-order alternatives (i.e. level of discontinuance or continuance). Note the interaction in the following figure.



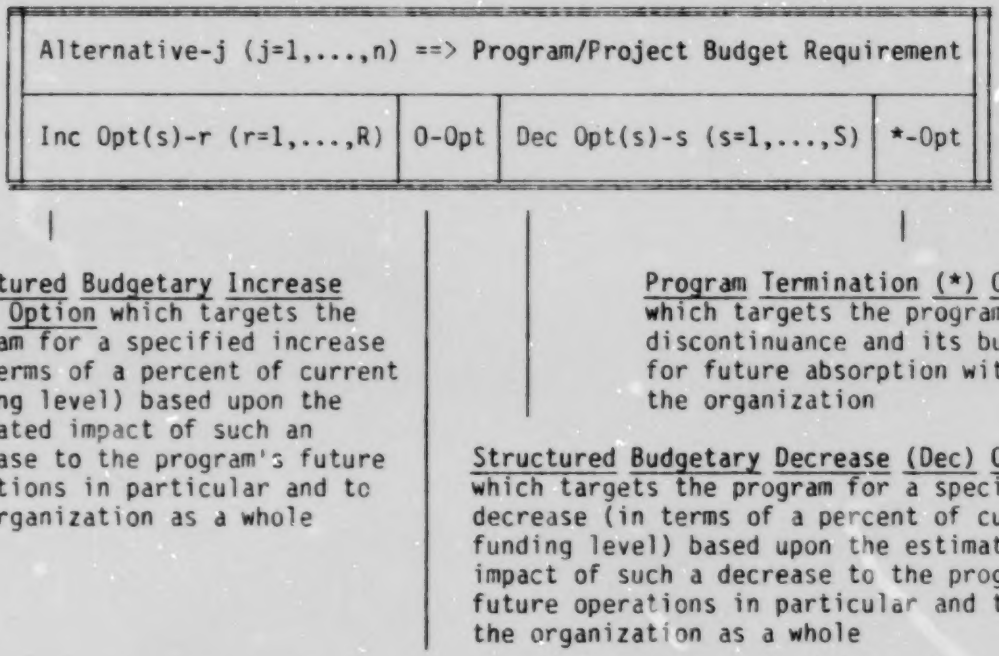
While each school is being considered for closure, the extent of that closure (sale, lease or mothball) is likewise being evaluated. At the same time, each school is being considered for operational continuance, and the level of that continuance. The higher-order level of decision alternative will automatically depend upon that alternative's lower-order (performance) feasibility. Indeed, the fate of one school in a highly transactional environment might prove a direct benefit or solemn toll for the fate of another school.

For a decision simulation model which evaluates varying levels of funding increase and/or decrease across all program budgets within an organization, the

BEST COPY

attributes of hierarchy, mutual exclusiveness and collective exhaustiveness must be augmented by a model which is also variable, rather than fixed, level. In the school closure model, each school had the same five lower-order alternatives available for evaluation. In the budget reduction model, however, each program budget under evaluation might have different options as lower-order alternatives. (See Reducing Higher Education Budgets Through Multiple Alternatives Modeling™ (1984) for a more detailed treatment of this subject.)

Consider the general form of the graduated budget reduction model:


Alternative-j (j=1,...,n) ==> Program/Project Budget Requirement
Inc Opt(s)-r (r=1,...,R)
0-Opt
Dec Opt(s)-s (s=1,...,S)
*-Opt
Structured Budgetary Increase (Inc) Option which targets the program for a specified increase (in terms of a percent of current funding level) based upon the estimated impact of such an increase to the program's future operations in particular and to the organization as a whole
Program Termination (*) Option which targets the program for discontinuance and its budget for future absorption within the organization
Structured Budgetary Decrease (Dec) Option which targets the program for a specified decrease (in terms of a percent of current funding level) based upon the estimated impact of such a decrease to the program's future operations in particular and to the organization as a whole
Budget Stasis or Zero (0) Option which targets the program for continued funding at the current level of budgetary requirement

such that:

(1) R (number/type of increases in funding options) $\geq$ 0;

(2) S (number/type of decreases in funding options) $\geq$ 0;

BEST COPY

(3) R not necessarily equal to S; and

(4) 0-Opt and *-Opt options considered as comprehensively reasonable across all programs.

Since each program budget may have different levels of funding increase and/or decrease available for evaluation, an illustrated sample might appear as:

Prog-01							Prog-02					Prog-03		...
+12	+03	0	-05	-12	-23	*	+07	+03	0	-15	*	0	*	...

<-------------------------------->
Program-01 has a wide variety of budgeting alternatives, including two options for increase and three options for decrease, in addition to the required alternatives of zero-budgeting (stasis) and complete termination of the program.

<---------------------->
Program-02 is more restricted than 01 in its available options, with two options for increase and only one avenue for decrease before the need for program termination.

<----------->
Program-03 is restricted to either funding at the current level, or the termination of the program.

Although the budget reduction model is variable-level hierarchical as compared to the fixed-level hierarchical characteristics of the school closure model, both models comply with the assumptions of mutual exclusiveness and collective exhaustiveness. Unfortunately, hierarchical models are not as straightforward in their validation procedures as are simple multiple alternatives frameworks. In addition, variable-level hierarchical models must be treated differently from fixed-level hierarchical models as we will observe in a later section.

The structure of the various decision alternatives (viz. courses of action) observed in the two aforegoing models was relatively simple and straightforward to construct. This is not always the case in decision modeling. Consider the case in which the decisionmaker is trying to decide which curriculum require-

BEST COPY

ments are best served by which specific software (courseware) packages on which particular microcomputer hardware systems.

For this modeling problem, the decisionmaker may wish to develop a rather convoluted hierarchical (probably variable-level) model which delineates each available hardware system by its compatible software packages, comparing each of the software packages across all instructional objectives and related classroom learning activities. For example:

Hardware Alternative-j (j=1,...,n) ==> Type of Microcomputer					
Software Alter-s (k=1,...,s,...,K)			Software Alter-t (k=1,...,t,...,K)		
Object 1.0	Object 2.0	Object 3.0	Object 1.0	Object 2.0	Object 3.0
Range of Possible Activities for 1.0	Range of Possible Activities for 2.0	Range of Possible Activities for 3.0	Range of Possible Activities for 1.0	Range of Possible Activities for 2.0	Range of Possible Activities for 3.0

Alternatively, the model builder may view the modeling context as exactly the opposite, delineating each instructional goal by its corresponding classroom learning activity, and then comparing available software and hardware for their matched suitability. For example:

BEST COPY

Learning Concept Alternative-j (j=1,...,n) ==> Instructional Goal																	
Classroom Activity #1						Classroom Activity #2						Classroom Activity #3					
Soft-A		Soft-B		Soft-C		Soft-A		Soft-B		Soft-C		Soft-A		Soft-B		Soft-C	
H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2	H1	H2

In actuality, neither of the above attempts at structuring the decision alternatives is more effective than simply viewing the modeling framework as a combination of two distinct, though not separate, models. Therefore, the modeling formulation displayed as:

Object 1.0			Object 2.0			Object 3.0		
A1	A2	A3	A1	A2	A3	A1	A2	A3

Hardware-A		Hardware-B		Hardware-C	
SW-1	SW-2	SW-1	SW-2	SW-1	SW-2

promotes a better, and more easily constructed alternatives framework for such a decisionmaking process. This formulation, in its currently fixed-level hierarchical state, can be easily reconstructed to represent a variable-level hierarchical model of the form:

Object 1.0			Object 2.0			Object 3.0		
A1	A2	A3	A4	A5	A6	A7	A8	A9

Hardware-A		Hardware-B		Hardware-C	
SW-1	SW-2	SW-1	SW-3	SW-4	SW-5

BEST COPY

where different activities for different instructional objectives are compared and contrasted for different software packages and instructional hardware systems.

In summary, the intended validation of any decision simulation model requires a clear understanding of the strengths and limitations fostered by the decisionmaker's choice of structure for the alternative courses of action to be evaluated. In general, the greater the structure, the easier and ultimately more reliable the validation of any solution identified via the model. From the structure of the alternatives, we must now turn our attention to the structure of the criteria which will be employed to value each alternative, and subsequently compare and constrast each alternative's worth as a potential member of the final solution (viz. decision) set.

CONTENT AND PROCESS CONSIDERATIONS IN DECISION MODELING

Validation of decisionmaking involves not only 'what' was utilized to compare and contrast the alternatives (i.e. the content or criteria) but also 'how' that criteria was used in the comparison and contrasting of those alternatives (i.e. the process or model itself). The decisionmaker approaches this validation of content and process by addressing two questions.

First, and as related to the criteria used to value each of the alternatives, validation seeks to answer:

> **To what extent did each of the criteria successfully compete for influence in the characteristics of the solution set? That is, does the decision (viz. each alternative chosen as solution) contain those criterion attributes desired by the decisionmaker?**

These questions are content oriented, related to the individual influence of each of the criteria used to value the alternative courses of action, and seek to identify which of the available criteria were ultimately responsible for the final solution (choice of alternatives).

BEST COPY

Secondly, the validation of the decisions derived from the model is concerned with questions related to process characteristics:

> To what extent did the model which compared and contrasted the alternatives adequately and appropriately determine a solution set from these alternatives based upon existing criteria? That is, does the simulation model demonstrate an unbiased process which (1) introduces criterion influence, (2) controls random selection, and (3) minimizes residual (criterion) influence?

The separateness between these two categories or phases of decision validation is not always clear at first reading. The first phase is concerned with the influence of the criteria, individually, on the final selection of decisional alternatives as solutions. The second phase is concerned with the collective influence of these criteria, i.e. confluence, as represented by the modeling framework which employed the criteria for comparing and contrasting the alternatives available for evaluation.

Some decision models simply compare alternatives by summing the numerical values (weighted or unweighted) of the criteria for each alternative, and then contrasting their composite sums. A number of serious problems are obviously associated with such an approach. The decisional framework, so hastily introduced in an earlier section of this paper, the Multiple Alternatives Model™, eliminates such problems by employing a system of simultaneous linear inequalities and equalities rather than a single sum of composite value. Before exploring the environment of criteria based modeling further, and its role in the eventual validation of criteria based decisions, let us first examine how numerical criteria compare and contrast alternative courses of action in two different settings.

Consider the situation where four alternatives are available for comparison and constrast, each alternative quantitatively measured (i.e. valued) on a scale of 1-5 across four criteria, where higher 'scores' represent more desirable effects. Such a modeling framework might be represented as follows:

BEST COPY

	Alter-1	Alter-2	Alter-3	Alter-4
Criterion-1	2	4	3	1
Criterion-2	1	4	5	5
Criterion-3	2	1	2	3
Criterion-4	1	1	1	1

An initial investigation demonstrates that no one alternative is highly desirable based upon its individual criterion measures. If a decisionmaker were to compute a sum for each alternative across all criteria, the alternatives be represented by the composite scores 6, 10, 11 and 10, respectively. While alternative-1 is certainly not a clear solution, alternative-2 and alternative-4 are likewise not similar in their criterion perspective although the composite sums are identical. If this model represented a singular-alternative situation, i.e. requiring one and only one alternative as a final solution, the choice of action would depend totally upon the weighted importance of the particular criteria. If the criteria were all weighted equally, an obviously dilemma is apparent. In the multiple alternatives setting, i.e. choosing as many actions as necessary to fit the needs of a problem, the decisionmaker would have to employ additional decisioning procedures or techniques to finally decide which alternatives were truly the 'best approach', collectively. Thus the problem with simple, composite summations when comparing and contrasting decisional alternatives.

Now contrast the above discussion with the following modeling parameters, wherein a multiple alternatives setting has been envoked, and where the model seeks to identify (simultaneously) the number of alternatives in a solution, and which ones (who) they are. The parameters state that any number of alternatives may be members of the solution if and only if their collective, individual sums across each criterion satisfy the associated limits on the right hand side of the model.

BEST COPY

	Alter-1	Alter-2	Alter-3	Alter-4	
Criterion-1	2	4	3	1	≥ 6
Criterion-2	1	4	5	5	≥ 7
Criterion-3	2	1	2	3	≤ 4
Criterion-4	1	1	1	1	≥ 2

A cursory review of this expanded model demonstrates that no single alternative could exist alone as the solution. Therefore, the multiple alternatives context is standard. A number of different combinations of alternatives satisfy one or more of the various criterion parameters, but as you have probably already concluded, only two alternatives collectively satisfy all of the criteria simultaneously -- alternatives 2 and 3. This example obviously exceeds the potential of the first example by not only detecting a solution, but by determining its number (how many) as well as type (who they are).

The remaining question, of course, lies in how to develop validation procedures which would effectively (adequately and appropriately) determine the true value of any decision(s) which were generated by such a model. Before attempting to deal more directly with an answer to this question, let us now examine two 'real' criteria based formulations for increased understanding.

Recall the mission of the hierarchical school closure model:

> TARGET THE NECESSARY NUMBER OF SCHOOL CAMPUSES IN THE SCHOOL DISTRICT FOR FUTURE OPERATIONAL DISCONTINUANCE OR CONTINUANCE AND THE SPECIFIC LEVEL OF THAT DIS/CONTINUANCE, RESPECTIVELY [choosing from three pre-evaluated options for closure (i.e. including sale, leasing or mothballing) and two options for operational continuance (functional partitioning with sub-site leasing and full program continuance (no change))]
>
> subject to:

BEST COPY

1. maximizing the positive influence from all school-based programs, and their support administrative and physical plant structures, on total school district functioning, based upon final school-by-school closure levels selected;

2. minimizing the negative influence from all school-based programs, and their support administrative and physical plant structures, on total school district functioning, based upon final school-by-school closure levels selected;

3. assuring that all schools and their programs within the school district are nondiscriminantly evaluated, compared and constrasted to ascertain their contributing value or worth to the goals of the school district as a whole; and,

4. conclusively targeting levels of operational discontinuance or continuance based upon valid measures of effectiveness, expenditure, economy, efficiency, and satisfaction pertinent to the character of each school, and its instructional mission.

To accomplish this intent, a reformulation of the multiple alternatives modeling™ framework (Wholeben, 1980) for hierarchical decisionmaking will be briefly illustrated. Within the MAM™ context, the operational alternatives for each school are delineated into five (5) closure/non-closure options depending upon a pre-modeling evaluation of each school, and the resulting school options scenarios or narratives which result. Criteria, in the form of numerical vectors, are introduced within the model to (1) simulate the decisional environment of the school district, its various service programs and instructional departments, (2) depict the strengths and limitations associated with each 'school closure' alternative, and (3) compete for influence in the comparison and constrast of each 'closure' alternative with other likely options.

A sample of the criteria portion of this approach to evaluating and selecting variant school closure alternatives based upon merit and estimated impact is MAM™ formulated as follows:

BEST COPY

	School Plant #1						School Plant #j						
	C/1	C/2	C/3	O/1	O/2	...	C/1	C/2	C/3	O/1	O/2		
Plant ADM Capacity				c_1	c_1	...				c_j	c_j	$\geq$	Futur Distr ADM
(expressed in students-count, from architectual blueprints)													
Plant ADA Vacancies				v_1	v_1	...				v_j	v_j	$\geq$	compd lower bound
(expressed in difference: plant capacity - plant ADM)													
School Prog Budgets				b_{14}	b_{15}	...				b_{j4}	b_{j5}	$\leq$	Distr Budg
(expressed in thousands of dollars, may be delineated by account)													
Anticipated Revenues	r_{11}	r_{12}		r_{14}		...	r_{j1}	r_{j2}		r_{j4}		$\geq$	compd lower bound
(expressed in thousands of dollars, may be delineated by source)													
Estimated Walk Dist	d_{11}	d_{12}	d_{13}	d_{14}	d_{15}	...	d_{j1}	d_{j2}	d_{j3}	d_{j4}	d_{j5}	$\leq$	compd upper bound
(expressed in average miles of site from attendance zones)													
Policy(s) Compliance	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	...	p_{j1}	p_{j2}	p_{j3}	p_{j4}	p_{j5}	$\geq$	compd lower bound
(expressed in frequency counts gathered through the scale below)													
	Fails to Meet Most Requirements [1]		Fails to Meet Some Requirements [2]		Marginally Acceptable (Neutral) [3]		Meets Most Requirements as Defined [4]		Exceeds Most Requirements as Defined [5]				
Opinion from Survey(s)	o_{11}	o_{12}	o_{13}	o_{14}	o_{15}	...	o_{j1}	o_{j2}	o_{j3}	o_{j4}	o_{j5}	$\geq$	compd lower bound
(expressed in frequency counts gathered through the scale below)													
	High Negative [1]		Negative [2]		More NEG Than POS [3]		More POS Than NEG [4]		Positive [5]		High Positive [6]		
Administrator Preference	a_1	a_1	a_1	a_1	a_1	...	a_j	a_j	a_j	a_j	a_j	$\geq$	compd lower bound
(expressed in frequencies, respondents indicating preferred sites)													
Control #1	1	1	1	1	1							=	1
Control #2							1	1	1	1	1	=	1

BEST COPY

Note how this model relies upon several types of measurement scales (nominal, ordinal and interval) as it attempts to quantitatively value each of the available decisional alternatives.

As a comparison, recall the mission of the hierarchical modeling framework for graduated budget reductions:

> ASSIGN FUNDING VARIANCES TO ALL BUDGETS WITHIN THE UNIVERSITY [choosing from pre-evaluated funding decrease and increase options, including stasis budgeting and the possibility of full program termination]
>
> subject to:
>
> 1. maximizing the positive influence from all programs on total university functioning, based upon final program-by-program budgeting levels selected;
>
> 2. minimizing the negative influence from all programs on total university functioning, based upon final program-by-program budgeting levels selected;
>
> 3. assuring all programs within the university are non-discriminantly evaluated, compared and constrasted to ascertain their contributing value or worth to the goals of the university; and,
>
> 4. conclusively targeting levels of funding based upon valid measures of effectiveness, efficiency, satisfaction and economy pertinent to the character of each individual college, school and/or department.

In this context, the sample criterion portion of the model might appear as:

BEST COPY

	Budget #1				Budget #2			Budget #3		
	00	15	25	**	00	05	**	00	**	
Personnel	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	<= $\$\$_1$
Example:	100	85	75	00	50	47	00	70	00	<= 195
	(personnel costs expressed in $1000 units)									
Personnel	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}	a_{17}	a_{18}	a_{19}	<= $\#\#_1$
Example:	6.0	4.5	4.0	0	4.0	3.5	0	3.0	0	<= 11.0
	(personnel requirements expressed in F.T.E. units)									
Policy 1.0	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{26}	a_{27}	a_{28}	a_{29}	>= b_2
Example:	4.7	3.6	2.3	1.9	2.5	2.1	1.3	3.9	3.8	>= 9.0

Fails to Meet Most Requirements	Fails to Meet Some Requirements	Marginally Acceptable (Neutral)	Meets Most Requirements as Defined	Exceeds Most Requirements as Defined
[1]	[2]	[3]	[4]	[5]

Policy 2.0	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{36}	a_{37}	a_{38}	a_{39}	>= b_3
Example:	2.4	2.3	1.5	1.1	4.9	4.5	2.3	3.6	1.0	>= 7.5
	(Scores assessed utilizing 5-point continuum displayed above)									
Opinion	a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	a_{46}	a_{47}	a_{48}	a_{49}	>= b_4
Example:	5.8	4.7	3.2	2.8	4.7	3.2	3.1	4.3	1.7	>= 16.0

High Negative	Negative	More NEG Than POS	More POS Than NEG	Positive	High Positive
[1]	[2]	[3]	[4]	[5]	[6]

Preference	a_{51}	a_{52}	a_{53}	a_{54}	a_{55}	a_{56}	a_{57}	a_{58}	a_{59}	>= b_5
Example:	25	50	10	15	66	33	1	85	15	>= 198
	(Data represent percents of frequency, where respondents indicated which of the three (3) program scenarios they perceived to be of benefit to the organization en total.)									
Control #1	1	1	1	1						= 1
Control #2					1	1	1			= 1
Control #3								1	1	= 1

BEST COPY

Note how the 'control submatrix' at the bottom of the figure assures that (1) each option within a particular program will be evaluated and (2) one and only one funding alternative will be selected for each such program.

This completes our introduction to the criteria based modeling of decision-making, and hopefully prepares the reader for a better understanding of the validation procedures which will now be examined in the remainder of this paper. We will start, now, with an examination of the various statistical procedures which prove helpful during validation, and then in later sections explore the validation process stage-by-stage.

STATISTICAL UTILITIES FOR DECISION MODEL VALIDATION

Validating criteria based decisions requires two separate evaluations. The first, an evaluation of criterion influence, seeks to establish the bivariate relationship between the alternatives which entered the solution set and those alternatives which did not, based upon differential measures of criterion impact between these two subsets. The second, an evaluation of criteria confluence, seeks to establish the multivariate relationship between these two alternatives subsets, the solution set and the non-solution set, based upon the collection of criterion variables which were available to the decisionmaking process for comparing and contrasting the original set of all alternatives. A third type of evaluation can also be instituted, that of simulation forecasting, which seeks to determine the extent to which the measured strength of the multivariate relationship might be utilized in later decisionmaking in the absence of further alternatives modeling.

Fortunately, standard parametric and nonparametric statistical procedures can be employed to facilitate all of the above three validation needs. The choice of statistics demands that (1) the information required for determining the extent of validation be pre-specified, and that (2) the scaling of both the dependent and independent variables be known. In our setting, the information required for determining validation is represented by the three evaluation needs: criterion influence, criteria confluence, and simulation forecasting.

BEST COPY

Variable scaling is restricted to the three conventional types of nominal, ordinal and interval. Unless otherwise noted, the term independent variable relates to measures of solution, whereas the term dependent variable relates to measures of the various criteria. Where appropriate, both the univariate case (viz. involving the analysis of only one criterion variable) and the multivariate case (viz. involving the analysis of several criteria as a group) will be addressed.

Obviously, validation for the single alternatives model, wherein only one alternative is selected (if any) as a decision, is trivial -- requiring only a survey of the various criterion values to compare the solution alternative with the remaining alternatives not selected. The multiple alternatives case, wherein several alternatives may be identified collectively as the solution, is not trivial but is relatively straightforward. Generally, the solution is displayed as a vector (i.e. a string of numerical values) of 1's and 0's, the 1's representing the alternatives which were selected as part of the solution, and the 0's as the alternatives which were not selected. Such dichotomous indicators when used as the values of a classification variable allow relative ease in the deployment of such statistics as oneway analysis of variance, contingency table analysis, and the Kruskal-Wallis analysis of variance procedure(s) for the univariate case, and the discriminant function analysis procedure for the multivariate case. We will briefly explore each of these procedures at this time.

The model validator is guided to the proper statistic to be used by understanding beforehand (1) what the validation is attempting to conclude, and (2) what the scaling(s) of the various dependent (criteria) and independent (solution) variables are. Criteria may be scaled as either nominal, ordinal or interval, and in the simple multiple alternatives modeling case the solution variable is a nominal (dichotomous 1-0) variable. We ignore direct address of the ratio case since the parameters for the interval case are identical. Our conclusions (directions) in the simple multiple alternatives case will address each of the issues of (1) criterion influence, (2) criteria confluence, and also (3) simulation forecasting. The context of criteria confluence receives greater attention in a later part of this section, when the decisionmaking environment might require several executions of the model under different parameters or contexts, arriving at a range of solution alternatives.

BEST COPY

Criterion influence is best served directly by analysis of variance procedures when the criterion variable (dependent variable) is either ordinal or interval. The parametric oneway analysis of variance procedure (or grouped t-test) is necessary when the criterion is intervally scaled, while the Kruskal-Wallis analysis of variance procedure (nonparametric, based upon rank and chi-squared formulations) is necessary when the criterion variable is ordinally scaled. Contingency table analysis (chi-squared analysis of the crosstabulation matrix) is necessary when the criterion variable is nominally scaled, but also recommended when an ordinally scaled criterion has a small number of ordinally based intervals. In the parametric sense, the grand distribution of criterion scores is segregated into two sub-distributions, one representing the distribution of those scores of the solution alternatives and the other representing the distribution of the scores of the non-solution alternatives. These distributions can then be compared to determine if:

(1) the means of the distributions are significantly different in the statistical sense (e.g. $p \leq .10$);

(2) the means of the distributions are significantly different in the magnitudinal sense (viz. the means of the distributions are different enough to portray important weight discrepancies between the solution and non-solution alternatives); and,

(3) the means of the solution-distributions of criteria are in the proper direction (viz. a positive benefit criterion distribution has higher values for the solution alternatives than for the non-solution alternatives, while a negative impact criterion distribution has lower values for the solution alternatives than for the non-solution alternatives).

The Kruskal-Wallis analysis of variance (nonparametric) technique provides a similar type of comparison strategy for ordinally scaled distributions which have three or more intervals involved, and for intervally scaled distributions where one or both of the assumptions of normality and homogeneity of variance

BEST COPY

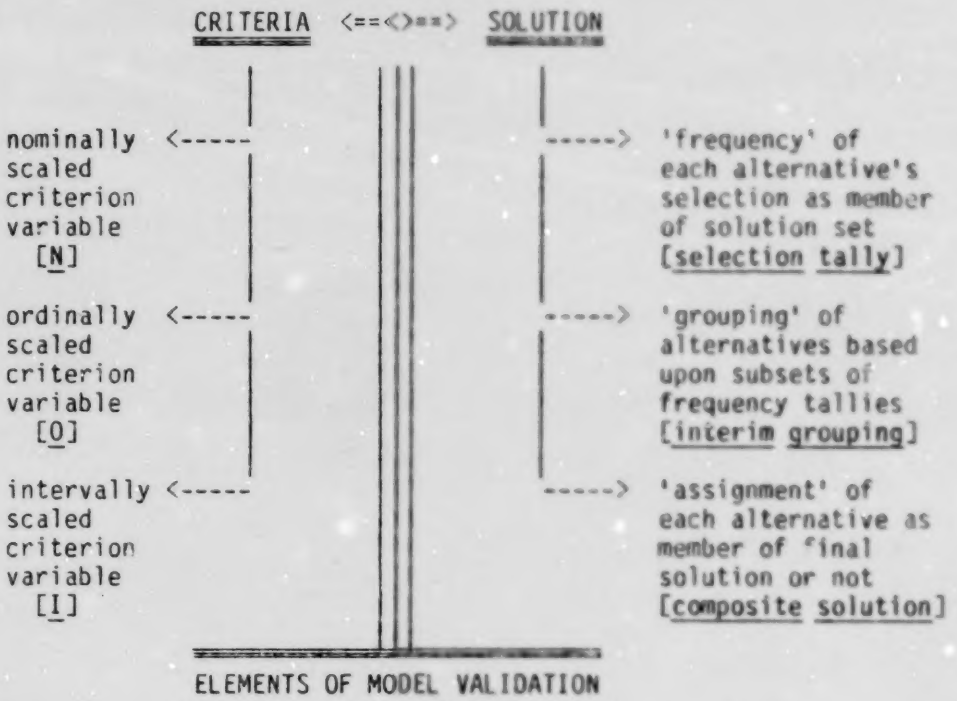
has been violated. Based upon the comparative ranks of the criterion scores, rather than metric measures of central tendency, the Kruskal-Wallis procedure employs the chi-squared distribution to test for the extent which composite rank differences between the solution and non-solution distributions are significantly different.

Contingency table analysis (i.e. the chi-squared formulation for crosstabulation matrices) must be employed for nominally scaled criteria, and for some ordinally scaled criteria. The test here, of course, is to determine the extent to which cell sizes differ across nominal categories of the criterion variable for the dichotomous categories (1-0) of the solution variable -- a difference which can be evaluated to depart significantly from what might have been expected from random occurrence. A significant chi-squared statistic, with cell weights in the desired direction(s), add to the validity of that criterion's influence in the solution.

In the multivariate case for simple multiple alternatives formulations, the validator seeks to measure the extent of criteria confluence within the solution or decision. This can only be achieved in this instance through the deployment of discriminant function techniques, a statistical procedure reserved normally for determining the simulation forecasting potential of the model. We will address this statistic momentarily.

The above discussion must be expanded significantly when our decisionmaking requirements include executing the model several times under different conditions. This is often the case because of the environmental characteristics of the operational system being highly fluid and transactional, and requiring the generation of solutions based upon dynamically changing environmental conditions. In the dynamic decisional environment, we must introduce the expanded approach to statistical validation presented in the following figure.

BEST COPY


CRITERIA <==<>==> SOLUTION
nominally scaled criterion variable [N]
ordinally scaled criterion variable [O]
intervally scaled criterion variable [I]
'frequency' of each alternative's selection as member of solution set [selection tally]
'grouping' of alternatives based upon subsets of frequency tallies [interim grouping]
'assignment' of each alternative as member of final solution or not [composite solution]
ELEMENTS OF MODEL VALIDATION

Each execution of a decisional model with multiple criteria across several different parameters (or constraints) has the potential of producing a different solution (collection of alternatives) as a result. The membership of these different solution sets (what alternatives are contained in the solution) and how this membership changes over the different executions are germane issues for evaluation by the decision validator.

Two new variables are introduced under these circumstances: the selection tally variable which displays the frequency with which each alternative was selected as part of several solutions over multiple executions, and the interim grouping variable which categorizes selection tallies by groups based upon their distribution of frequencies. In this case, of course, the selection tally variable is intervally scaled, and the the interim grouping variable ordinally scaled. The dichotomous 1-0 solution variable mentioned previously in this section will now be referred to as the composite solution variable, and represents our final decision based upon the characteristics of the selection tally and interim grouping variables. Our use of statistics can now be formulated through the use of the following tables.

BEST COPY

For the case in which the criterion variable is nominally scaled:

CRITERIA			SOLUTION		
nominal variable	ordinal variable	interval variable	selection tally	interim grouping	composite solution
[M]	[O]	[I]	[I]	[O]	[N]

X----------X
Oneway analysis of variance or grouped t-test is the preferred technique in the univariate case, with Kruskal-Wallis analysis of variance or contingency table analysis as alternate in situations of low-number alternative models. Separate analyses are required for the multivariate case.

X----------X
Kruskal-Wallis (nonparametric) analysis of variance is the preferred technique in the univariate case, with contingency table analysis as alternate. Separate analyses are required for the multivariate case.

X----------X
Contingency table analysis is the preferred technique in the univariate case. Separate analyses are required for the multivariate case.

BEST COPY

For the case in which the criterion variable is ordinally scaled:

CRITERIA			SOLUTION		
nominal variable	ordinal variable	interval variable	selection tally	interim grouping	composite solution
[N]	[O]	[I]	[I]	[O]	[N]

X-------------------------X
Oneway analysis of variance is the preferred technique in the univariate case when the number of ordinal intervals is small, with Kendall's tau correlation coefficient as alternate. Separate analyses are required for the multivariate case.

X-----------------------------------X
Both contingency table analysis and Kendall's tau correlation coefficient are the preferred technique in the univariate case, with Kruskal-Wallis analysis of variance as alternate when the number of ordinal intervals is large. Separate analyses are required for the multivariate case.

X---X
Contingency table analysis is the preferred technique in the univariate case when the number of ordinal intervals is small, with Kruskal-Wallis analysis of variance as alternate. Separate analyses are required for the multivariate case.

BEST COPY

Finally, for the case wherein the criterion variable is intervally scaled:

CRITERIA			SOLUTION		
nominal variable	ordinal variable	interval variable	selection tally	interim grouping	composite solution
[N]	[O]	[I]	[I]	[O]	[N]

X----------X
Simple regression is the preferred technique in the univariate case. Stepwise multiple regression can be employed in the multivariate case.

X----------X
Oneway analysis of variance is the preferred technique in the univariate case when the ordinal levels are small, with Kendall's tau correlation coefficient as alternate. Separate analyses are required for the multivariate case.

X----------X
Oneway analysis of variance is the preferred technique in the univariate case. Discriminant function analysis can be employed for the multivariate case.

Criterion influence is still evaluated by comparing each individual solution with each of the criterion variables as previously discussed in an earlier part of this section. Criterion confluence is now more directly evaluated through the use of the selection tally and interim grouping variables. Through the use of these two new variables, the inner workings of the criteria as a group in addition to the utility of the model as a criteria comparison process can both be evaluated for their validity.

BEST COPY

VALIDATION STAGE I: SINGULAR CRITERION INFLUENCE

By introducing the reader (hopefully somewhat successfully, in the previous section) to the types and deployments of the several statistical procedures recommended for final decision validation, when decisionmakers seek to evaluate potential alternative solutions for their individual contribution to an overall final decision, it was necessary to briefly address the three validation goals of criterion influence, criteria confluence and simulation forecasting. In this section, and the two following, we will expand upon this discussion in greater detail.

The assessment of criterion influence is singular in nature, as contrasted with criteria confluence, and seeks to measure the extent of influence of each criterion in the final solution. For example, if a criterion was utilized in a decision model which sought to:

> "close those schools with high maintenance costs as measured over the past five years, costs measured in thousands of dollars",

the decisionmaker would expect the characteristics of the final solution to demonstrate that the solution alternatives (schools to close) would have higher maintenance costs associated with their history than did non-solution alternatives. Alternatively, if a budget reduction model sought to:

> "target those instructional programs for graduated funding reductions which demonstrate higher enrollment declines over the past seven years, enrollments measured in head counts",

the decisionmaker would likewise hope for targeted reductions in those programs which have lower enrollment trends over time as compared to the non-solution alternatives.

Singular criterion influence can be assessed in both the simple multiple alternatives case (one solution is generated) and the cyclic multiple alter-

BEST COPY

natives case (several executions of the model for a changing environment which generate several solutions). The deployment of statistics is identical for both the simple and cyclic multiple alternative cases.

The validation of a decision seeks to determine the extent to which pre-stated goals of the solution model are ultimately satisfied by the decision. These goals are usually stated in terms of criteria which act as constaints to the decisional process. Criterion based goals, however, are often in conflict with each other due to the competition which normally exists between the criteria, themselves. For example, consider the two goals:

1. to identify and select at least two instructional courseware packages which will operate on a Brand-X computer system, and will provide for at least 80 percent satisfaction of pre-stated classroom activity requirements; and,

2. to identify and select at least one, but not more than three, classroom activities for each instructional objective based upon each of the concept learning goals required under current state-level mandates.

These two goals can be easily delineated into objective-based statements which can then directly be input to the decision model as individual criterion constructs:

1.1 two or more instructional courseware packages must be selected;

1.2 all selected instructional courseware package must have be ability to execute within the Brand-X operating system;

1.3 all instructional courseware packages selected must collectively address 80 percent or more of the activity requirements of the classroom;

2.1 each of the conceptual learning goals required under state-level mandates must be satisfied by the final solution generated;

2.2 each of the instructional objectives which underlie the state-level conceptual learning goals must be satisfied by the final solution generated;

BEST COPY

2.3 three or less, but at least one, classroom activity must be selected to fulfill the instructional objectives of learning.

From these six constructs, criteria (in our case, numerical) must be identified which satisfactorily address each of the construct's intentions to constrain the formation of the final decision (solution alternatives). As it appears initially, nearly all of the above constructs can be easily (relatively speaking, of course) formulated as individual criteria by the use of frequency vectors, nominally scaled, and consisting of 0's and 1's. The reader is once again directed to the bibliography for further detailed study of this concept. The exception is objective (construct) 1.3 : all instructional courseware packages selected must collectively address 80 percent or more of the activity requirements of the classroom;

For construct 1.3, a determination of (1) how many criteria will be required to model the construct effectively (adequately and appropriately) and efficiently, and (2) what the scaling of each such criterion will be, must first be accomplished. This is not trivial. Since courseware packages may duplicate many of the activity requirements, some technique for denoting distinct courseware contribution must be achieved. Secondly, the relationship between the activities satisfied, and the courseware as the 'satisfier', must also be evaluated. Finally, some method for signifying the 'percent of satisfaction' between the activities and the courseware must be accomplished.

To accomplish the mission of validation, the decisionmaker must validate the means by which the model builder constructed the model for the decisionmaking process, including the foci of the various constructs utilized to guide and constrain the comparison and contrast of potential solution alternatives. This is not a statistical process, but rather requires the validator to survey the reasoning, rationale and interpretation of each criterion in support of the particular construct to which it contributes its influence.

BEST COPY

VALIDATION STAGE II: MULTIPLE CRITERIA CONFLUENCE

Contrasted with singular criterion influence, the assessment of criteria confluence is concerned with measuring the collective influence of the criteria (as a group) input to the model. Statistically, such processes demand multivariate approaches where possible, and separate analyses consolidated into a single table or figure where multivariate techniques are precluded due to the scaling of the dependent (criteria) variables. Typically, only intervally scaled criterion variables can be directly analyzed in the multivariate statistical sense. All other variables must be interpreted for their collective confluence based upon consolidated separate analyses.

Assessing the confluence resident within a group of criteria, and the contribution of such confluence to overall solution and decision validity, expands significantly the information obtained from singular criterion influence analyses. In effect, the assessment of confluence seeks to better interpret the competition which existed between and among the total set of criteria as referenced in the criterion distributions of the solution alternatives subset, and as compared to the non-solution subset of alternatives. By allowing all criteria to relate simultaneously (in the multivariate statistical sense) to an independent solution variable (viz. the selection tally, interim grouping and/or composite solution variables), the decisionmaker understands which of the criteria are more influencial in the solution achieved, the degree or weight of that influence, and therefore which of the criteria are not as influencial as might originally had been desired. This post-hoc assessment of competition between and among criteria, based upon the solution as an independent variable, addresses the issues of preference and trade-off in solution formulation.

Preference typically refers to those relationships between criteria and solutions which are desirable based upon the model builder's original intentions and decisional requirements. Trade-off, on the other hand, refers to those criteria-solution relationships which are other-than-desirable (although not always so negative as to be undesirable) based upon the modeler's decisional requirements. Preference and trade-off evaluations are as much a part of singular criterion influence assessments as they are of multiple criteria confluence

BEST COPY

evaluations. Basically, preference is significantly noteworthy if and only if all of the following conditions are met:

1. the means, medians or mean-ranks of the measured criterion for the solution-alternatives sub-distribution is in compliance with pre-stated decisional requirements (e.g. a goal-based model which seeks to close schools with high safety risk, based upon student-walking requirements, identifies a group of schools whose distributional characteristics on the safety criterion demonstrate higher safety risk factors than schools not selected for closure);

2. the magnitudinal differences between these means, medians or mean-ranks are interpreted by the decisionmaker as significantly large enough to warrant importance (e.g. a goal-based model which seeks to identify prospective teachers for a formal interview, based in part on their number of years living in geographic regions considered highly multi-cultural, identifies a group of teacher interviewees whose distributional characteristics on this multi-cultural environment criterion demonstrate a mean of 15 years, contrasted to a mean of 3 years for teachers who were not selected for interview); and,

3. the differences in central tendency between the solution and non-solution distributions are interpreted as statistically significant (e.g. $p \leq .10$).

In a later section, criteria confluence will be illustrated in two contexts, the first a summary of assessments of singular criterion influence, and the second a direct assessment via the use of discriminant function analysis.

VALIDATION STAGE III: SIMULATION FORECASTING

In all evaluation and simulation modeling, of course, these is an element of forecasting -- the prediction of likely effects or impact based upon the alternatives and competitive criteria available to the model. If a model is designed to target levels of funding increase on the basis of zero-base budgeting, it is assumed that the funding levels selected will allow those programs to impact the organization as positively as possible, given the constraints of the environment. Therefore, the model forecasts maximum positive benefit and minimum negative side-effects based upon the solution generated.

BEST COPY

Another form of prediction which is often possible during decision validation, is that of a direct assessment of a model's ability to forecast through the use of its simulation processes. This third, and final, stage of validation is typically executed through the use of discriminant function analysis -- a parametric, multivariate statistical procedure which uses the solution variable as an independent variable initially, and then as a dependent variable during what is referred to as the 'reclassification stage'. Promised several times earlier within as many sections, we will explore discriminant function analysis in an attempt to demonstrate how this procedure validates a model's ability to perform simulation forecasting.

Like multiple regression, discriminant function analysis attempts to establish the relationship between a single independent variable, called the classification variable because it is usually nominally scaled, and multiple dependent variables which must be intervally scaled, or at least defined as dummy variables. The procedures develop discriminant functions (equations which discriminate between levels of the classification variable, one for each ordinal interval), entering each criterion stepwise based upon the strength resident in its distinct and unique influence to discriminate between the levels of the classification variable (our solution variable). Once the ability to discriminate has been optimized, that is, the entry of any further criteria is declared as inconsequential for an increased ability to discriminate (or until all criteria have been entered into the procedure), the discriminant function(s) then attempts to use the criteria as independent variables to predict the classification membership in the original classification variable for each case in the data base. The success of this prediction (i.e. its accuracy) is the final test of the discriminant function, called the reclassification stage.

If the decisionmaker interprets that:

(1) the alternatives heretofore modeled will not change substantially for the immediate future;

BEST COPY

(2) the criterion variables entered into the construction of the discriminant function(s) are particularly important to the decisionmaking environment when compared to other criteria available to the model;

(3) the criterion variables entered are themselves highly transactional and indicative of a rather fluid or dynamic decisional environment; and,

(4) the reclassification accuracy of the discriminant function(s) is acceptable for strategic planning purposes, ...

then the discriminant function(s) can be used in the future for producing 'what if' agenda without the need to remodel the decisional environment in a multiple alternatives setting discussed previously. By simply entering new criterion values into the discriminant function(s), new information as to which alternatives seem plausible based upon this data can be generated. This is the attribute of simulation forecasting.

In those evaluation models where a decision actually eliminates one or more alternatives from further consideration (e.g. closing a school), the use of a discriminant function(s) for forecasting purposes must be careful to include only those alternatives which are viable. Once an alternative is eliminated from any further (future) consideration, the format of the discriminant function must also change to effectively note this new environment.

TOOLS AND MECHANICS OF MAM™ DECISION VALIDATION

With our introduction to decision validation now hopefully complete, our attention will be directed towards the specific case of the cyclic model. The cyclic model is one in which several executions are performed under different conditions to observe any subsequent changes in the multiple alternatives solution which result. This is the preferred and recommended process to follow in most, if not all, decisional modeling episodes. Such multiple executions also allow the development of those solution variables known as the selection tally and interim grouping variables, in addition to the composite solution variable, and can direct a more rationale construction of the composite solution variable if deployed correctly and consistently.

BEST COPY

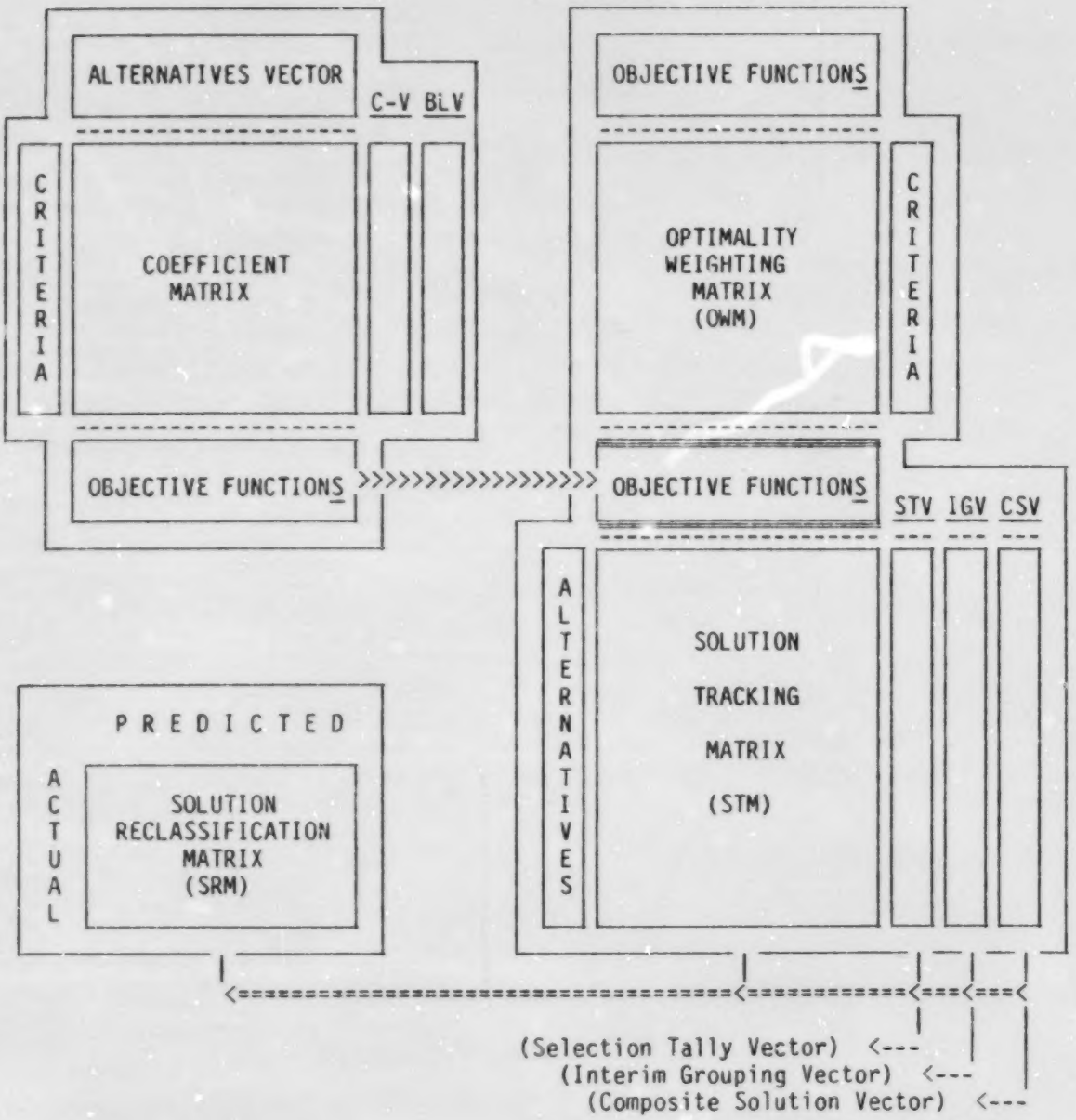
Let us review the process of cyclic modeling. A model is developed which guides the comparison and contrast of multiple alternatives by means of multiple competition-inducing criterion variables. Since the environment of decision-making is usually both dynamic and transactional -- constantly changing, and effecting other parts of the environment with these changes -- simple alterations to the model can produce variable solutions based upon these alterations. Although the assessment and validation of cyclic modeling proceeds in identical fashion to those techniques previously discussed in this paper, such assessment can be significantly facilitated mechanically by the use of a number of information organization tools. In this section, we will introduce the reader to these tools, their general description and function, and then in the following sections illustrate their function by actual numerically based examples.

Mechanics suggest a structure for dealing with a specific process compleat with a schedule of procedureal steps which must be accomplished in a recommended order. Tools suggest that specific formats for organizing information and results are recommended in specific frameworks already designed -- a fill-in-the-blanks arrangement, if you will. For our purposes of decision validation, we will use the mechanics and tools developed for the Multiple Alternatives Model™ (Wholeben, 1982) to provide a framework for validation of the cyclic modeling and decision generation process.

The tools of decision modeling and solution validation can be displayed as follows:

BEST COPY

MULTIPLE ALTERNATIVES MODEL™ (MAM™)



BEST COPY

For the purposes of this paper, the term 'objective functions' represents the varying parameters of our model based upon cyclic re-executions. The reader need not be concerned with how this is actually accomplished in the MAM™ framework to enjoy (?) the remainder of this section.

Our initial concern is with that part of the figure labelled the 'solution tracking matrix', or STM. The STM provides an organizational structure for tracking the various solutions which result from the varying parameters of the model's parameters, that is, attempting to discern any continuity in the solutions for a dynamic and transactional environment. Typically, there is one column in the STM for each of the executions performed. The rows of the STM are the alternatives being compared, and the cell entries for the STM are a '1' if an alternative was selected as a solution for that particular environment parameter (column of the STM), a '0' otherwise. Therefore, ten executions produce a STM with ten columns of 1's and 0's.

The next tool to examine is the STV, or selection tally vector, which we have previously discussed in a preceding section. The STV is an intervally scaled frequency variable which simply stores the row-wise sums of the STM. Thus if alternative-6 (row 6 of the STM) held three 1's (selected three times as a solution), the sixth row of the STV would have the value '3'.

To the right of the STV is the interim grouping vector, or IGV, also introduced earlier in this paper. The IGV is an ordinally scaled variable which was constructed by interpreting the distributional characteristics of the STV. Consider the following example:

BEST COPY

STV	IGV
6	3
3	2
8	3
7	3
4	2
8	3
0	1
1	1
4	2
7	3

which (quite) arbitrarily constructed an IGV from the sample STV based upon the following distribution of the STV:

```
                  x        x  x
x  x     x  x     x     x  x  x
-  -  -  -  -  -  -  -  -  -  -
0  1  2  3  4  5  6  7  8
```

where 'natural' breaks appear to exist between the ranges 0-1, 3-4, and 6-8. In the case of a STV where the STM represents a large number of alternatives (rows) the model validator might employ standard deviations from the mean to categorize the STV into an IGV. Regardless of the construction technique employed, the IGV should be considered ordinal rather than simply nominal since the classification categories constructed represent increasing frequency sums found in the STV. This also expands the opportunities for validation through statistical means.

The CSV, or composite solution vector, stores only 1's and 0's, and represents our final decision based upon all previous mechanics, tools and analysis results.

Above the solution tracking matrix, lies the optimality weighting matrix. The OWM stores the results of the analysis of variance tests which seek to determine the extent of singular criterion influence in the various alternative solutions stored in the STM. Similar tools can be constructed for tests which require contingency table analyses because of the criterion variable's nominal (or ordinal) scaling, and for the Kruskal-Wallis procedures which are used with

BEST COPY

ordinally scaled criterion variables. Notice how the OWM is concerned with matching criterion influence (rows defined as criteria) to the 'objective functions', while the STM cross-matches alternatives selected as solutions with the 'objective functions' which produced such solutions.

The last figure in the above diagram is that of the 'solution reclassification matrix', which is concerned with displaying the results of the discriminant function analysis procedures -- always for the CSV, usually for the IGV, and occasionally even for the columns of the STM. More on this in a later section.

THE SOLUTION TRACKING MATRIX

The solution tracking matrix, or STM, is only utilized in the multiple alternatives case when cyclic modeling is present. Recall that cyclic modeling is employed whenever several alternate decisions are necessary based upon some modification of the original criteria based model. The STM 'tracks' the separate solutions (1-0 solution variables) which result from each of these executions and re-executions.

Although not physically part of the matrix, but closely related to it, are the three variables: the selection tally vector, interim grouping vector, and the composite solution vector. Each of these three vectors are directly dependent upon the internal 1-0 configuration of the STM.

The structure of the solution tracking matrix, and its peripheral and subsidiary elements, may be physically viewed as follows:

BEST COPY

SOLUTION TRACKING MATRIX

	(Objectives Functions Used)					STV	IGV	CSV
	---	---	---	---	---			
Alternative-01	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
Alternative-02	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
Alternative-03	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
Alternative-04	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
Alternative-05	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
Alternative-06	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
Alternative-07	0-1	0-1	0-1	0-1	0-1	0-j	0-k	0-1
SLV								

(Solution Limit Vector)

j = number of alternatives
k = number of decisions
(such that $0 \leq k \leq j$)

SOLUTION LIMIT MATRIX

Optimality Criteria	N of Cyclic Optimizations				
Obj Func 01					
Obj Func 02					
Obj Func 03					
Obj Func 04					
Obj Func 05					

STV =:
Selection Tally Vector

IGV =:
Interim Grouping Vector

CSV =:
Composite Solution Vector

BEST COPY

Although it is beyond the scope of this paper, the additional elements of the solution limit vector (variable) and solution limit matrix are displayed in the figure as well. These two subsidiaries are used when the decision environment requires a step-wise solution generation approach, allowing the modeler to better witness the model's approach to equilibrium -- wherein no further decision is necessary since all criterion objectives have been satisfied.

A sample solution tracking matrix might then appear as follows:

BEST COPY

Individ Budgets	< Varying Parameters for Dynamic Decisionmaking > CR-0	CR-1	CR-2	CR-3	CR-4	CR-5	CR-6	CR-7	CR-8	CR-9		SELECT TALLY	INTERIM GROUP
01	X	X	X	X	X	X	X	X	X	X	===>	10	4
02	X	X	X	X	X	X		X			===>	7	3
03	X	X	X				X	X			===>	5	2
04	X	X	X	X	X	X	X	X	X	X	===>	10	4
05	X		X	X	X	X	X		X	X	===>	8	3
06	X			X	X		X	X	X		===>	6	2
07	X	X		X	X	X	X	X	X	X	===>	9	3
08											===>	--	1
09	X	X	X	X	X	X	X	X	X	X	===>	10	4
10						X					===>	1	1
11		X	X	X		X		X		X	===>	6	2
12											===>	--	1
13					X		X				===>	2	1
14									X		===>	1	1
15	X	X	X	X	X	X	X	X	X	X	===>	10	4
16						X		X	X	X	===>	4	2
17	X	X	X	X	X	X	X	X	X	X	===>	10	4
18				X					X		===>	2	1
19		X									===>	1	1
20											===>	--	1
21	X		X								===>	2	1
22								X			===>	1	1
23		X	X		X		X		X	X	===>	6	2
24											===>	--	1
25											===>	--	1
26	X	X					X			X	===>	4	2
27				X							===>	1	1
28				X					X		===>	2	1
29										X	===>	1	1
30						X					===>	1	1
31											===>	--	1
	12	12	11	13	11	12	12	12	13	12		(Solution Set N)	
	-208	-208	-194	-200	-209	-209	-200	-218	-218	-193		(Model Cuts)	
	685	685	699	693	684	684	693	675	675	700		(New Budget Amt)	

Note: Total Initial Budget = 893.5 ($1000's)

BEST COPY

Each 'X' represents the selection of a particular alternative (row) as a member of the cyclic solution set for each of the separate executions of the model based upon changing parameters of a dynamic and transactional environment (columns). Column totals signify: (1) the number of alternatives required to comprise the solution set, (2) how much money was reduced in total from all available budgets, and (3) the projected budget requirements under the simulated system. Row totals provide a summary of how many times each particular alternative was chosen as a solution, and a sample interim grouping vector based upon the following STV distribution:

```
<-- 1 -->       <-- 2 -->   <-- 3 -->  <4>

    x
x   x
x   x                                    x
x   x   x                                x
x   x   x                    x           x
x   x   x        x           x           x
x   x   x        x   x   x   x   x   x   x
--- --- --- --- --- --- --- --- --- --- ---
000 001 002 003 004 005 006 007 008 009 010
```

In summary, the mission of the solution tracking matrix is to: (1) summarize the solution tracking patterns for each of the cyclic re-executions, and (2) provide easy access to the first tool of validation, the optimality weighting matrix.

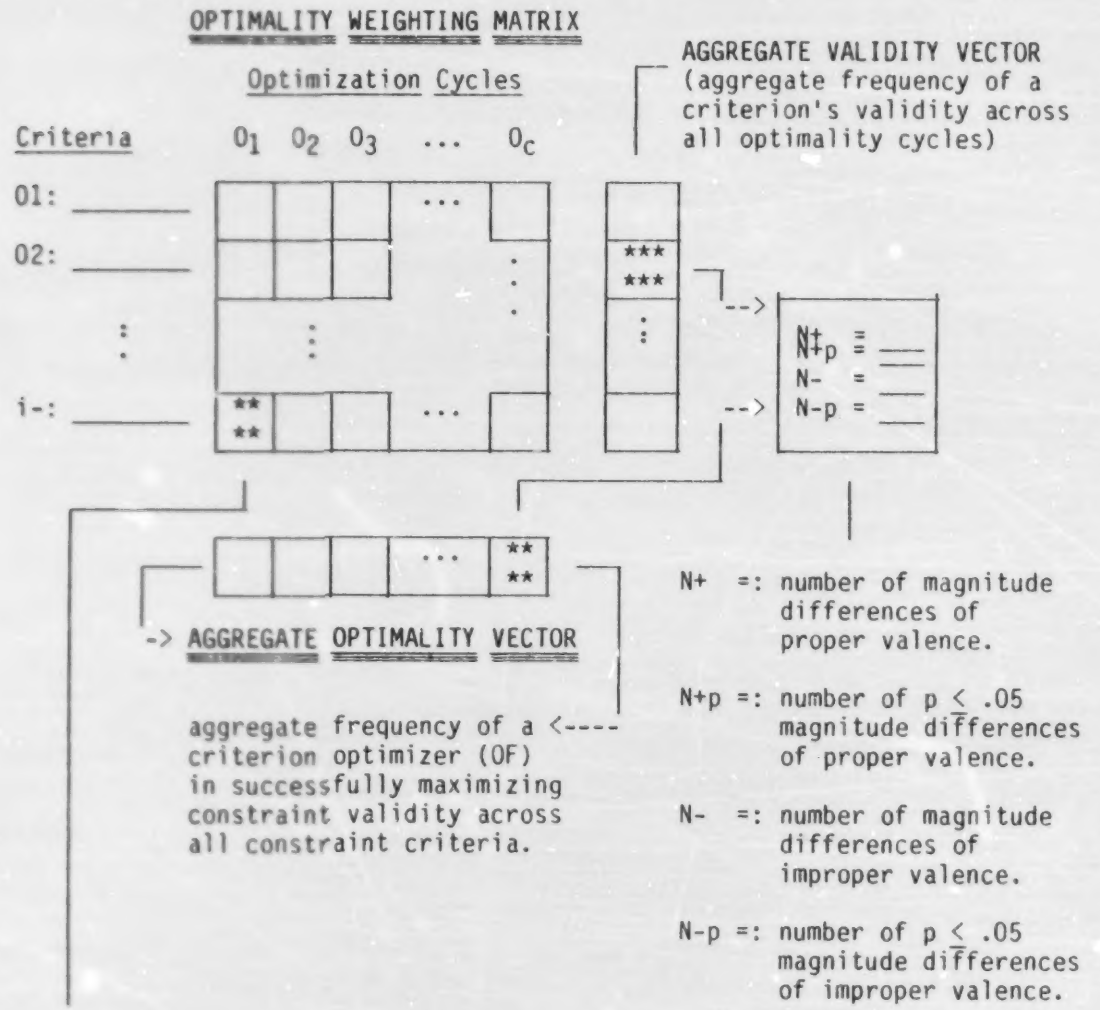
THE OPTIMALITY WEIGHTING MATRIX

The optimality weighting matrix, or OWM, summarizes the extent to which a cyclic decisional model exemplifies the pre-stated objectives (i.e. criteria) of the model builder. Therefore, the OWM is concerned directly with the discussion of criterion influence, and indirectly with the issue of criteria confluence. For the purposes of this paper, we will illustrate the OWM framework which is used when intervally scaled criteria are analyzed by analysis of variance (parametric) procedures.

BEST COPY

The optimality weighting matrix may be conceptualized as follows:

BEST COPY


OPTIMALITY WEIGHTING MATRIX
Optimization Cycles
Criteria
O_1 O_2 O_3 ... O_C
01:
02:
i-:
AGGREGATE VALIDITY VECTOR
(aggregate frequency of a criterion's validity across all optimality cycles)
N+p =
N- =
N-p =
-> AGGREGATE OPTIMALITY VECTOR
aggregate frequency of a criterion optimizer (OF) in successfully maximizing constraint validity across all constraint criteria.
N+ =: number of magnitude differences of proper valence.
N+p =: number of p ≤ .05 magnitude differences of proper valence.
N- =: number of magnitude differences of improper valence.
N-p =: number of p ≤ .05 magnitude differences of improper valence.

$MN_{(Select)}$ = ____ ----------> criterion distribution mean for the decision subset;

$SD_{(Select)}$ = ____ ----------> criterion distribution standard deviation for decision subset.

$MN_{(No/Select)}$ = ____ ----------> criterion distribution mean for the non-decision subset;

$SD_{(No/Select)}$ = ____ ----------> criterion distribution standard deviation for non-decision subset.

$P_{(Select\ v.\ N/Select)}$ = ____ ---> significance probability level between the decision subsets.

BEST COPY

To better understand how the optimality weighting matrix actually proceeds in its contribution to the determination of decision validation, we will employ a actual field-based example. In a later section, the direct address of criteria confluence will be accomplished through an extension of this example.

To illustrate the use of the optimality weighting matrix, and its ability to display both criterion influence and criteria confluence resulting from a series of analysis of variance procedures with an independent variable as the dichotomous 1-0 solution variable and all intervally scaled criterion variables as dependent variables, we present the following case study.

A school board has decided to validate the budget reduction decision(s) of the school district budget officer by attempting to discover which of the stated criteria in fact support the decisions formally recommended, and which do not. Taking the reductions (alternatives) offerred as a solution, and the criteria utilized to generate these reductions, the school board generates appropriate analysis of variance statistics by comparing each criterion with the solution. Since a series of ten (10) solutions were formally presented to the board, based upon ten separate executions of the decision model (cyclic modeling in a dynamic and transactional environment), nine additional series of analysis of variance procedures had to be conducted. The results of these statistical analyses are displayed below, where the rows represent the criteria of the model, and the columns represent the parameters which were altered for each separate execution.

In this model, the decisionmaker (district budget officer) sought to maximize the effect of each criterion within the generated solution(s) of the model(s).

BEST COPY

(Objective = MAXIMIZATION)

	C-01	C-02	C-03	C-04	C-05	C-06	C-07	C-08	C-09	C-10	$N_c(p\leq.10)$	
CRIT-01	.00	.17	.08	.19	.05	.14	.02	.14	.24	.20	=P	
	40.4	27.3	30.2	26.4	31.3	24.2	32.4	28.0	25.8	26.8	=I	[04]
	6.1	14.4	13.4	14.3	12.9	16.3	11.2	13.9	14.7	14.7	=X	
CRIT-02	.28	.00	.01	.86	.47	.67	.23	.23	.59	07	=P	
	22.1	35.5	30.4	17.1	20.5	18.5	22.8	22.9	13.3	26.0	=I	[03]
	12.3	3.9	8.3	15.4	13.7	14.6	11.9	11.9	18.2	9.9	=X	
CRIT-03	.48	.09	.00	.81	.71	.97	.87	.52	.47	.32	=P	
	16.7	21.7	28.7	11.8	15.1	12.8	13.9	16.4	9.5	18.2	=I	[02]
	10.7	7.5	4.4	13.9	11.8	13.1	12.5	10.8	15.6	9.7	=X	
CRIT-04	.05	.14	.05	.00	.02	.00	.22	.04	.02	.12	=P	
	34.4	31.5	35.0	47.4	37.5	40.1	30.1	34.9	35.5	31.8	=I	[07]
	15.3	17.2	16.0	4.9	14.6	11.7	18.1	15.0	13.5	17.0	=X	
CRIT-05	.02	.12	.05	.05	.00	.16	.00	.16	.02	.08	=P	
	30.6	26.5	29.7	28.2	43.3	26.2	36.2	25.8	29.6	27.7	=I	[07]
	9.7	12.3	11.2	10.2	3.7	12.4	6.2	12.7	9.2	11.5	=X	
CRIT-06	.76	.86	.76	.81	.99	.00	.32	.28	.96	.58	=P	
	20.9	21.7	24.7	24.1	22.8	39.8	16.8	29.2	22.5	26.1	=I	[01]
	23.9	23.4	21.6	21.7	22.7	11.9	26.5	18.6	22.9	20.6	=X	
CRIT-07	.75	.75	.67	.82	.05	.75	.07	.75	.82	.75	=P	
	4.2	4.2	4.5	3.8	9.1	4.2	8.3	4.2	3.8	4.2	=I	[02]
	2.6	2.6	2.5	2.8	0.0	2.6	0.0	2.6	2.8	2.6	=X	
CRIT-08	.00	.00	.00	.04	.02	.00	.01	.09	.05	.01	=P	
	51.0	45.5	47.8	40.5	43.2	48.5	43.4	55.0	39.6	44.4	=I	[10]
	15.2	18.7	18.8	20.8	21.4	16.8	20.1	12.8	21.5	19.5	=X	
CRIT-09	.31	.73	.59	.02	.83	.61	.86	.82	.06	.56	=P	
	49.8	49.2	51.3	45.2	49.5	48.9	49.6	49.5	46.2	48.7	=I	[00]
	50.2	50.5	49.3	53.5	50.3	50.7	50.2	50.3	52.8	50.8	=X	
CRIT-10	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	=P	
	57.8	57.1	63.6	53.3	62.2	57.0	57.8	56.3	52.0	58.3	=I	[10]
	10.5	10.9	9.7	11.1	13.7	11.0	10.5	11.5	12.1	10.2	=X	
$N_r(p\leq.10)$	[05]	[04]	[07]	[05]	[06]	[04]	[05]	[03]	[05]	[04]		

BEST COPY

where the triple-entry cells (P, I and X) hold the values:

P = the statistical significance of mean differences, include vs. exclude;
I = mean budget amount for included object of budget revision; and,
X = mean budget amount for excluded object of budget revision.

Both row-wise and column-wise summations are also available. The column-wise summation, $N_c(p \leq .10)$, represents the frequency of $p \leq .10$ occurrences, where criterion constraint values reflect desireable mean-value weights across the cyclic objective functions. The row-wise summation, $N_r(p \leq .10)$, likewise represents the frequency of $p \leq .10$ occurrences, where each defined objective function produced desireable mean value weights across the criterion constraints.

From this optimality weighting matrix, we can ascertain that the first execution of the model (the first solution, the first column) produced a solution which demonstrated a statistically significant difference in criterion means between the solution and non-solution distributions ($N_r(p \leq .10)$) for a total of five (5) times, while the third execution produced seven (7) such desirable effects.

Surveying the $N_c(p \leq .10)$ summations, we might note that criteria 8 and criteria 10 produced significant distributional mean differences in each of the 10 solutions generated, despite the dynamic modifications introduced within the model over these ten executions. Criterion 6 produced only one such difference, for execution 6.

The optimality weighting matrix summarizes information relating to criterion influence and criteria confluence. For example, we can see that the solution generated during the initial execution of the model produced an criterion-1 effect where the criterion-1 value of the solution measures 40.4 whereas the non-solution measures 6.1. This suggests that the goal of maximizing criterion effect was achieved in this instance, and therefore the decision of the budget officer supported by OWM evidence. At the same time, however, distributional comparisons for criterion-2 in the first execution suggest that, although the

BEST COPY

criterion value is larger for the solution alternatives than for their non-solution counterparts, the difference is not significantly different statistically, meaning that some alternatives of higher criterion-2 value were not chosen as solutions while some alternatives of lower value were. These are all important notations for determining the validity of the solution.

In the discussion of criterion-1 (above) we are tabulating the occurrences of analysis results which speak in favor of the decision or solution. Such tabulations are known as preference factors; that is, we preferred that criterion-1 be significantly greater, statistically, during the first execution for solution alternatives than for non-solution alternatives, and we achieved the preferred result. However, in the discussion of criterion-2 we found a result which was not as we had preferred, but which occurred due to the model's need to find a solution which matched all competitive and contrasted criteria, and had to sacrifice some elements of criterion-2 to accomplish this end. The ability to satisfy decisional requirements while sacrificing some elements of certain criteria is called 'satisficing', and introduces the validation need to examine all trade-off factors in the modeling formulation.

To demonstrate this concept more effectively, only the more important preference and trade-off factors (in the previous OWM) will be displayed for convenience. The reader is cautioned, however, that the blank cells in the following figure also represent trade-offs or satisficing, but are not as important as the ones we will now discuss.

BEST COPY

(Objective = MAXIMIZATION)

	C-01	C-02	C-03	C-04	C-05	C-06	C-07	C-08	C-09	C-10	$N_c(p \leq .10)$	
CRIT-01	.00 40.4 6.1		.08 30.2 13.4		.05 31.3 12.9		.02 32.4 11.2				=P =I =X	[04]
CRIT-02		.00 35.5 3.9	.01 30.4 8.3						.59 13.3 18.2	07 26.0 9.9	=P =I =X	[03]
CRIT-03		.09 21.7 7.5	.00 28.7 4.4	.81 11.8 13.9		.97 12.8 13.1			.47 9.5 15.6		=P =I =X	[02]
CRIT-04	.05 34.4 15.3		.05 35.0 16.0	.00 47.4 4.9	.02 37.5 14.6	.00 40.1 11.7		.04 34.9 15.0	.02 35.5 13.5		=P =I =X	[07]
CRIT-05	.02 30.6 9.7		.05 29.7 11.2	.05 28.2 10.2	.00 43.3 3.7		.00 36.2 6.2		.02 29.6 9.2	.08 27.7 11.5	=P =I =X	[07]
CRIT-06	.76 20.9 23.9	.86 21.7 23.4				.00 39.8 11.9	.32 16.8 26.5		.96 22.5 22.9		=P =I =X	[01]
CRIT-07					.05 9.1 0.0		.07 8.3 0.0				=P =I =X	[02]
CRIT-08	.00 51.0 15.2	.00 45.5 18.7	.00 47.8 18.8	.04 40.5 20.8	.02 43.2 21.4	.00 48.5 16.8	.01 43.4 20.1	.00 55.0 12.8	.05 39.6 21.5	.01 44.4 19.5	=P =I =X	[10]
CRIT-09	.91 49.8 50.2	.73 49.2 50.5		.02 45.2 53.5	.83 49.5 50.3	.61 48.9 50.7	.86 49.6 50.2	.82 49.5 50.3	.06 46.2 52.8	.56 48.7 50.8	=P =I =X	[00]
CRIT-10	.00 57.8 10.5	.00 57.1 10.9	.00 63.6 9.7	.00 53.3 11.1	.00 62.2 13.7	.00 57.0 11.0	.00 57.8 10.5	.00 56.3 11.5	.00 52.0 12.1	.00 58.3 10.2	=P =I =X	[10]
$N_r(p \leq .10)$	[05]	[04]	[07]	[05]	[06]	[04]	[05]	[03]	[05]	[04]		

BEST COPY

In the above preference/trade-off sub-matrix of the OWM, 37 percent of the cells are blank, demonstrating that 37 percent of the available criterion influence and confluence was 'satisficed'. This form of satisficing is known as positive trade-off, since the solution set means were still larger than the non-solution set means. As we will demonstrate later, there is a type of satisficing known as negative trade-off, where the decision displays diametrically opposed results based upon criterion influence.

We now further delineate the OWM sub-matrix into what is called the preference factor submatrix, that is, displaying only those cells where we can make definitive judgements regarding criterion objectives which were both adequately and appropriately satisfied according to pre-stated, desired effects. The extent to which a solution matches the criterion-referenced objectives of the decisionmaking episode must be significant in terms of the percentage of cells in the OWM containing statistically significant preference factors.

Observe the following preference factor submatrix:

BEST COPY

(Objective = MAXIMIZATION)

	C-01	C-02	C-03	C-04	C-05	C-06	C-07	C-08	C-09	C-10	$N_c(p \leq .10)$	
CRIT-01	.00 40.4 6.1		.08 30.2 13.4		.05 31.3 12.9		.02 32.4 11.2				=P =I =X	[04]
CRIT-02		.00 35.5 3.9	.01 30.4 8.3							.07 26.0 9.9	=P =I =X	[03]
CRIT-03		.09 21.7 7.5	.00 28.7 4.4								=P =I =X	[02]
CRIT-04	.05 34.4 15.3		.05 35.0 16.0	.00 47.4 4.9	.02 37.5 14.6	.00 40.1 11.7		.04 34.9 15.0	.02 35.5 13.5		=P =I =X	[07]
CRIT-05	.02 30.6 9.7		.05 29.7 11.2	.05 28.2 10.2	.00 43.3 3.7		.00 36.2 6.2		.02 29.6 9.2	.08 27.7 11.5	=P =I =X	[07]
CRIT-06						.00 39.8 11.9					=P =I =X	[01]
CRIT-07					.05 9.1 0.0		.07 8.3 0.0				=P =I =X	[02]
CRIT-08	.00 51.0 15.2	.00 45.5 18.7	.00 47.8 18.8	.04 40.5 20.8	.02 43.2 21.4	.00 48.5 16.8	.01 43.4 20.1	.00 55.0 12.8	.05 39.6 21.5	.01 44.4 19.5	=P =I =X	[10]
CRIT-09											=P =I =X	[00]
CRIT-10	.00 57.8 10.5	.00 57.1 10.9	.00 63.6 9.7	.00 53.3 11.1	.00 62.2 13.7	.00 57.0 11.0	.00 57.8 10.5	.00 56.3 11.5	.00 52.0 12.1	.00 58.3 10.2	=P =I =X	[10]
$N_r(p \leq .10)$	[05]	[04]	[07]	[05]	[06]	[04]	[05]	[03]	[05]	[04]		

BEST COPY

The preference factor submatrix of the OWM demonstrates 46 percent compliance with the pre-stated objectives of criterion influence within the decisional model. Note that all preference-factor cells must address our earlier discussed demand for model validation: statistically based mean differences in the appropriate magnitudinal direction.

Once the extent of both positive satisficing and preference factor has been summarized, validation turns to the issue of negative satisficing -- those elements of the solution which were necessary in order to find a solution matching all the competitive influences existing within the model's multiple criteria. The existence of negative satisficing does not immediately invalidate a model, but the extent of such factor(s) should be made aware to the individuals who must ultimately accept or reject the solution offerred, in our case, the school board.

The extent of negative satisficing, or trade-off, is displayed below in a further delineation of our original OWM. Only those cells which display that the mean of the non-solution alternatives distribution is larger than the the mean of the solution alternatives distribution are tabulated as negative satisficers. Recall that the school district budget officer sought to maximize the impact of each criterion in the final solution.

This trade-off sub-matrix is shown as:

BEST COPY

(Objective = MAXIMIZATION)

	C-01	C-02	C-03	C-04	C-05	C-06	C-07	C-08	C-09	C-10	
CRIT-01											=P =I =X
CRIT-02									.59 13.3 18.2		=P =I =X
CRIT-03				.81 11.8 13.9		.97 12.8 13.1			.47 9.5 15.6		=P =I =X
CRIT-04											=P =I =X
CRIT-05											=P =I =X
CRIT-06	.76 20.9 23.9	.86 21.7 23.4					.32 16.8 26.5		.96 22.5 22.9		=P =I =X
CRIT-07											=P =I =X
CRIT-08											=P =I =X
CRIT-09	.91 49.8 50.2	.73 49.2 50.5		.02 45.2 53.5	.83 49.5 50.3	.61 48.9 50.7	.86 49.6 50.2	.82 49.5 50.3	.06 46.2 52.8	.56 48.7 50.8	=P =I =X
CRIT-10											=P =I =X

BEST COPY

In the above negative-satisficing, trade-off submatrix, a total of 17 percent of the cells contain criterion influence unacceptable (relatively) to the original intent of the model. Of especial interest are the two situations in which criterion-09 has significantly lower values statistically for the solution than for the non-solution, a direct contradiction of pre-stated objectives.

In summary, our validation of criterion influence and confluence allows the following statements:

1. preference factor index of 46 -- meaning that 46 percent of the cells in the optimality weighting matrix represented compliance with modeling objectives which were statistically significant;

2. positive satisficing index of 37 -- meaning that 37 percent of the cells in the optimality weighting matrix represented non-statistically significant compliance with pre-stated modeling objectives; and,

3. negative satisficing index of 17 -- meaning that 17 percent of the cells in the optimality weighting matrix were not in compliance with pre-stated modeling objectives.

With the above validation information, the school board could reasonably accept or reject the solution offerred by the school district budget officer. Obviously, additional information regarding the weighted importance of the various criteria, the weighted importance of the different dynamic and transactional parameters introduced during successive executions of the model, and other demands of the district not included for analysis in the model, must also be incorporated within this final decisionmaking effort.

THE SOLUTION RECLASSIFICATION MATRIX

We have seen, that with careful interpretation of the optimality weighting matrix, not only direct measures of criterion influence can be associated with separate and distinct solutions but also an indirect understanding of criteria confluence is discernible from an overview of the OWM. In this section, we will explore a technique which will address more directly this issue of criteria confluence, and its appropriateness for validating decisionmaking.

BEST COPY

Criteria confluence represents the collective or competitive influences existing between and among criteria when multiple alternatives are to be compared and contrasted for solution membership. Where criterion influence allows an interpretation of separate criterion-solution relationship(s), criteria confluence seeks answers to the questions:

Which criteria are most representative of the solution generated?

and of these representative criteria,

What are their relative strengths in discriminating between the solution and non-solution alternatives?

To answer these two questions, we resort to the use of the statistical procedure, discriminant function analysis, and the MAM™ validation tool, the solution reclassification matrix, or SRM.

The reader will recall from a previous section that discriminant function analysis is a regressive technique which constructs discriminant function(s) to relate a nominally scaled classification variable to ranges and combinations of measures associated with several intervally scaled criterion variables. The discriminant fuction is thus a mix between multiple regression and factor analytical procedures. As a stepwise regression approach, the strongest criterion (viz. the criterion with the most influence, relatively, considering its competitive status with the remaining criteria) is entered as an dependent variable in a discriminant function, differentiating between the levels of the classification variable (i.e. the solution) as much as possible. Then with the remaining residual variance left unexplained, a second criterion is entered into the discriminant function, itself the strongest discriminator based upon the other criteria which have not been entered, and additional variance is explained. This process continues until the variables are all used, or until no further variance control is possible.

BEST COPY

For example, consider the following solution summary which might have been available from the solution tracking matrix:

ALTER	FREQ	ALTER	FREQ	ALTER	FREQ	ALTER	FREQ
01	7	09	5	17	7	25	0
02	5	10	0	18	4	26	4
03	3	11	5	19	0	27	0
04	7	12	0	20	0	28	1
05	5	13	1	21	1	29	0
06	4	14	0	22	1	30	0
07	6	15	7	23	3	31	2
08	0	16	5	24	0		

where 31 alternatives have been modeled across 10 criteria, for a total of ten cyclic executions, and the above selection tallies (i.e. FREQ) computed.

For purposes of discussion, let us assume that we are continuing to validate the school district budget officers funding recommendations, that the alternatives are distinct budget reduction proposals by program, and that the criteria used to evaluate these program budgets are standard account categories. Using the selection tally vector as the classification (solution) variable, and the criteria as a multivariate collection of discriminating variables, discriminant function analysis might have produced the following stepwise results.

STEP	ENTERED	REMOVED	NOT USED
1	Budgetary Composites		Certificated Salaries
2	Employee Benefits		Classified Salaries
3	Contractual Services		Instructional Supplies
4	Travel Expenditures		Administrative Perception
5	Capital Outlay		
6	Supplies and Materials		

Here we witness that six (6) of the ten (10) criteria are found to discriminate between the levels of the STV, with the remaining four (4) criteria unable

BEST COPY

to explain any additional variance in the classification variable. We also can see that the criterion 'budgetary composites' was the strongest of the discriminating independent variables, and 'supplies and materials' the weakest.

Our next objective is to compute the extent to which these six criteria accurately discriminate between the levels of the STV. To accomplish this, discriminant function analysis attempts to re-predict into which classification level each of the alternatives falls, and then compare the actual classification with the predicted classification. Such results might appear as follows:

Actual Group	(N)	0	1	2	3	4	5	6	7
0	(11)	9:82	1:10	1:09	--	--	--	--	--
1	(4)	1:25	3:76	--	--	--	--	--	--
2	(1)	--	--	1:100	--	--	--	--	--
3	(2)	--	--	--	2:100	--	--	--	--
4	(3)	--	--	--	--	3:100	--	--	--
5	(5)	--	--	--	--	1:20	3:60	1:20	--
6	(1)	--	--	--	--	--	--	1:100	--
7	(4)	--	--	--	--	--	--	--	4:100

which represents an average percent of 84 percent of the grouped cases accurately re-classified. Coupled with the information summarized in the optimality weighting matrix, measures of both criterion influence and criteria confluence are now available to the school board, in order to make a final determination as to accept the budget officer's recommendations, or not.

The structure of the solution reclassification matrix may thus be viewed conceptually as follows:

BEST COPY

Predicted Solution Membership of Program Budget Scenarios

	Funding at Level 'x'	Not Funded
Funding at Level 'x' Crisis Budgeting Model Solution, where Select Budgets are funded at Allocation Level 'x'.	CONTRIBUTION TO SOLUTION RELIABILITY frequency and percent of budget scenarios correctly reclassified	CONTRIBUTION TO SOLUTION UNRELIABILITY frequency and percent of budget scenarios not correctly reclassified
Not Funded Remaining Budgeting Model Solution, where Scenarios are not funded at any level.	CONTRIBUTION TO SOLUTION UNRELIABILITY frequency and percent of n/budget scenarios not correctly reclass/	CONTRIBUTION TO SOLUTION RELIABILITY frequency and percent of n/budget scenarios correctly reclassified

Solutions correctly reclassified contribute to the measure of criterion confluence, while inaccurate reclassifications do not. However, and especially in the case of a multi-level classification variable, the degree to which an inaccurate reclassification occurs is also important. Consider the following expanded figure of the solution reclassification matrix.

BEST COPY

Predicted Solution-Weights Membership of Budget Scenarios

Weights Assigned at	Wgt=0	Wgt=1	Wgt=2	Wgt=3	Wgt=4	Wgt=5	Wgt=6
'0'	REL	U/REL	---------->				
'1'	U/REL	REL	---------->				
'2'	<---	U/REL	REL	U/REL	---------->		
'3'	<----------		U/REL	REL	U/REL	---------->	
'4'	<----------			U/REL	REL	U/REL	--->
'5'	<----------				U/REL	REL	U/REL
'6'	<----------					U/REL	REL

In this figure, the degree of inaccurate reclassification, i.e. the extent to which the re-predicted level is deviant (left or right) from the actual classification level, can be studied for its importance; that is:

<- extent solution under-weighted <- REL -> extent solution over-weighted ->

This also allows somewhat of a 'marginal error' computation, in which the decisionmaker might allow a ±1 error factor in computing the final accuracy of reclassification.

With the conceptual framework of the solution reclassification matrix in mind, and the potential for a computation of marginal error, let use view the SRM in action with another example.

BEST COPY

BUDGET	FREQ	BUDGET	FREQ	BUDGET	FREQ	BUDGET	FREQ
01	10	09	10	17	10	25	0
02	7	10	1	18	2	26	4
03	5	11	6	19	1	27	1
04	10	12	0	20	0	28	2
05	8	13	2	21	2	29	1
06	6	14	1	22	1	30	1
07	9	15	10	23	6	31	0
08	0	16	4	24	0		

SUMMARY OF CRITERION VALUE IN DISCRIMINATING SELECTION FREQUENCIES:

STEP	ENTERED	REMOVED	NOT USED
1	Budgetary Composites		Certificated Salaries
2	Instructional Supplies		Employee Benefits
3	Employee Benefits		Supplies and Materials
4	Capital Outlay		Contractual Services
5	Administrative Perception		Travel Expenditures
6	Classified Salaries		
7		Employee Benefits	

CLASSIFICATION RESULTS FROM PREDICTIVE VALIDATION:

Actual Group	(N)	0	1	2	4	5	6	7	8	9	10
0	(6)	4:67	1:17	1:17	--	--	--	--	--	--	--
1	(7)	1:14	5:71	1:14	--	--	--	--	--	--	--
2	(4)	1:25	1:25	2:50	--	--	--	--	--	--	--
4	(2)	--	--	--	2:100	--	--	--	--	--	--
5	(1)	--	--	--	--	1:100	--	--	--	--	--
6	(3)	--	--	--	--	--	2:67	--	--	1:33	--
7	(1)	--	--	--	--	--	--	1:100	--	--	--
8	(1)	--	--	--	--	--	--	--	1:100	--	--
9	(1)	--	--	--	--	--	--	--	--	1:100	--
10	(5)	--	--	--	--	--	--	1:20	1:20	--	3:60

Percent of grouped cases correctly classified: 70.97

NOTE: Accepting a tolerance of ±1 weight (tolerance factor = 9% (1/11))
Percent of corrected grouped cases correctly classified: 88.50

BEST COPY

In this example, an expanded number of references to the alternatives are found in the selection tally vector, and more interaction results in the stepwise regression procedure(s). Reclassification of the solution is greatly subdued as compared with our earlier illustration (71 percent compared with 84 percent), but with a ± 1 marginal error, an adjusted accuracy of 89 percent is calculated. Again, all highly usable information for the members of the school board.

Discriminant function analysis and the solution reclassification matrix can also be utilized to evaluate the composite solution (1-0) vector as is evidenced in the following figure.

BEST COPY

BUDGET	INCL	BUDGET	INCL	BUDGET	INCL	BUDGET	INCL
01	1	09	0	17	1	25	0
02	1	10	0	18	0	26	0
03	0	11	1	19	0	27	0
04	1	12	0	20	0	28	0
05	1	13	0	21	0	29	0
06	0	14	0	22	0	30	0
07	1	15	1	28	1	31	0
08	0	16	1	24	0		

SUMMARY OF CRITERION VALUE IN DISCRIMINATING INCLUSION DECISIONS:

STEP	ENTERED	REMOVED	NOT USED
1	Budgetary Composites		Certificated Salaries
2	Contractual Services		Classified Salaries
3	Instructional Supplies		Employee Benefits
			Supplies and Materials
			Travel Expenditures
			Capital Outlay
			Administrative-Perception

CLASSIFICATION RESULTS FROM PREDICTIVE VALIDATION:

(Percents in Parenthesis)

Actual Group	(N)	Predictive Group Membership 0	1
0	21	21 (100.00)	--
1	10	1 (10.0)	9 (90.0)

Percent of Grouped Cases Correctly Classified: 96.77

BEST COPY

As levels in the classification vector decrease, which is usually the case when the selection tally vector (or interim grouping vector) are transformed into the composite solution vector, reclassification accuracy will increase (see above). This is additional rationale for employing the SRM with both the STV and IGV as well as the CSV, and observing a conservative approach to decision validation.

SPECIAL CONSIDERATIONS IN HIERARCHICAL MODEL VALIDATION

The validation of a hierarchical model poses a special problem, not insurmountable but tedious, for the decisionmaker. The problem lies in the structure of the alternatives vector (actually its substructure), and in the fact that all hierarchical models force preclusion of a certain subset of alternatives once a particular member of that subset has been selected as the solution member.

For example, once a school has been determined to remain operational at its present level, the sub-alternatives of closure with either sale, lease or mothball options are necessarily precluded from further consideration. In the budget reduction model, if the program warrants termination then an alternative of a 15 percent increase in funding is no longer relevant.

The hierarchical model has, as you might remember, one of two types of levels associated with its structure -- fixed or variable. The fixed-level hierarchical model is illustrated by the school closure model, since the same five alternatives are available to each school site. The budget reduction model, on the other hand, is variable-level, with different options available to each program based upon likely-estimate scenarios.

As you might guess, the fixed-level hierarchical model provides the greatest opportunity for several varying approaches to decision validation. For example, schools targeted for closure and sale can be compared with schools targeted for closure and lease, or with schools targeted for functional partition and sub-site lease. Schools functionally partitioned and sub-site leased can be compared with schools remaining unchanged in operation, and so forth. The solution validator need only remember to structure the data base wherein resides the criterion data set. As the experienced evaluator will attest, even this task of managing the criterion data base can be tedious in itself.

BEST COPY

The variable-level hierarchical model must be carefully understood as one approaches the validation of any decisions generated by its execution(s). Four general types of decisions are available: graduated funding decrease, graduated funding increase, current funding level, and total program termination. Since the graduated options (increase and/or decrease) will in all likelihood have several levels themselves, validation possibilities increase geometrically. For example, in comparing terminated programs with programs remaining operational, contrasting evaluations must exist for all levels of operational programs, including those program also decreased in their funding allotments.

While the statistics remain the same, the validator must resort to many more analyses of variance, for example, in order to construct an optimality weighting matrix which is functionally informative -- the OWM itself also increasing in size, and therefore in complexity.

SUMMARY AND RECOMMENDATIONS

In this paper, we have explored the theoretical and conceptual contexts in which decision evaluation and validation reside. We have explored the use of statistical procedures for their role in decision validation, and introduced a number of tools and procedures which will assist the decisionmaker in the validation of a decision model, its processes and results. We have also, hopefully, made clear the need for employing validation procedures in all criterion based decisionmaking and solution generation from available alternatives.

Finally, we hope that the reader, if not already a believer in data based evaluation and decisionmaking, now sees the opportunity for employing such techniques in the future for many of the complex issues facing educators today.

BEST COPY

[Selected Readings and References]

An Advanced Multiple Alternatives Modeling Formulation for Determining Graduated Fiscal Support Strategies for Operational and Planned Educational Programs. Presented to the 1984 Annual Meeting of the American Educational Research Association, SIG on Fiscal Issues, Policy and Educational Finance, New Orleans; April 24, 1984.

Spread Sheet Simulation as an Adjunct Training Device for Fiscal Policy Analysis Instruction. (with Ronald A. Lindahl) In "Fifth Annual Convention Proceedings for the National Educational Computing Conference" (pp. 235-241) Iowa City: University of Iowa, 1983.

Demography, Economics and Rationalism: The Future of Higher Education. INVITED ADDRESS to the 1983 Convocation of Northwest Regional Associates of the Danforth Foundation; Union, Washington; April 1983.

MICROPIK: A Multiple-Alternatives, Criterion-Referenced Decisioning Model for Evaluating CAI Software and Microcomputer Hardware Against Selected Curriculum Instructional Objectives. (No. 73 in the series, "Research on Evaluation Program") Portland, Oregon: Northwest Regional Educational Laboratory, June 1982, 123 pp.

Multiple Alternatives for Educational Evaluation and Decision-Making. (No. 72 in the series, "Research on Evaluation Program") Portland, Oregon: Northwest Regional Educational Laboratory, May 1982, 416 pp.

Operational Network Displays. In "Communication Strategies in Evaluation" (N. Smith, Ed.) Beverly Hills, California: SAGE Publications, Inc., 1982, 320 pp.

The Modeling of Policy Alternatives. Issue paper within the AERA symposium: "Policy Alternatives: The Management of Decline". Accepted symposium at the 1982 American Educational Research Association Annual Meeting, New York City, March 20, 1982.

Discriminant Function Analysis as a Post-Hoc Operations Research Determinant of Criterion Strength and Binary Decisioning Reliability. Accepted paper to the Special Interest Group on Multiple Linear Regression Applications, 1982 American Educational Research Association Annual Meeting, New York City, March 20, 1982.

BEST COPY

A Computerized, Operations Research Model for Matching Microcomputer Hardware Available Software and Desired Instructional Objectives. INVITED PAPER to Special Interest Group on Computer Applications to Instruction, 1982 American Educational research Association Annual Meeting, New York City, March 20, 1982.

Needs Assessment for Changing Communities and Resources: Computer Simulation Models. INVITED PAPER to the Washington Educational Research Association Conference, "Effective Needs Assessment: Meeting the Challenge", Seattle, Washington, March 17, 1982.

Techniques in Cost/Benefit Analysis for Implementing the Use of Computers Educational Research Association, Seattle, Washington (May 1981).

Multiple Alternatives Modeling in Determining Fiscal Roll-Backs During Educational Funding Crises. (with J. Sullivan) (No. 70 in the series, "Research on Evaluation Program") Portland, Oregon: Northwest Regional Educational Laboratory, November 198., 126 pp.

The Design, Implementation and Evaluation of Mathematical Modeling Procedures for Decisioning Among Educational Alternatives. Lanham, Maryland: University Press of America, 1980. (453 pp.)

How to Determine Which School to Close. NASSP Bulletin: National Association of Secondary School Principals, 1980, 64:39, 7-12.

Curriculum Activity Packaging. Paper presented to the Washington Educational Research Association Conference, Seattle, Washington (April 1980).

Evaluation of Elementary School Sites for Targeted Closure. Madison Metropolitan School District, Madison, Wisconsin. April 1979.

BEST COPY